

Univerzita Karlova v Praze

Filozofická fakulta

Ústav českého jazyka a teorie komunikace

Bakalářská práce

Jan Henyš

Krácení ve dvojslabičných českých apelativech deklinačního typu *žena* (korpusový výzkum)

Vowel length variation in two-syllable nouns of the *žena* declension
type (corpus research)

Praha 2016

Vedoucí práce: doc. Mgr. Robert Adam, Ph.D.

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

Praha 14. 8. 2016

.....
Jan Henyš

Abstrakt

Tato práce se zabývá alternací kořenného vokálu, která se projevuje u některých dvojslabičných feminin deklinačního typu *žena*. První část práce představuje přehled zachycení tohoto jevu ve vybraných mluvnicích češtiny, vydáních Pravidel českého pravopisu a dalších publikacích. V druhé části je v jednotlivých fázích popsána extrakce alternujících dvojslabičných feminin z českého synchronního korpusu psané češtiny. Z revidovaných výsledků je pak vytvořen seznam všech alternujících dvojslabičných feminin deklinačního typu *žena*, který je založený na korpusových datech. Na základě dat získaných z korpusu jsou následně ověřována vybraná tvrzení z literatury uvedená v první části práce, která se týkají např. vlivu lexikálního významu nebo hláskové stavby jednotlivých slov na alternaci kořenného vokálu.

Klíčová slova: skloňování, substantivum, vzor *žena*, kvantita, alternace, korpusový výzkum

Abstract

This thesis aims on vowel variation in two-syllable nouns of the *žena* declension type. The first part of this work focuses on ways of capturing the matter in chosen grammars of Czech language, various editions of *Pravidla českého pravopisu* (*The Rules of Czech Orthography*) and other publications. The second part contains a description of particular stages of extracting two-syllable nouns of the *žena* declension type with vowel alternation present. From the revised results is then created a list of all the nouns containing varying vowel based solely on the data obtained from corpus. This part also aims on verification of various statements from Czech grammars or other publications, which are relating e.g. to the influence of lexical semantics or to phonetic structure of words on the vowel variation.

Keywords: declension, noun, declension type *žena*, vowel length, alternation, corpus research

Obsah

Úvod	5
1 Krácení kořenného vokálu v dvojslabičných apelativech deklinačního typu <i>žena</i>	6
1.1 Zachycení jevu v literatuře	6
1.1.1 Česká mluvnice Jana S. Tomíčka	6
1.1.2 Brus jazyka českého	7
1.1.3 Trávníčkova Mluvnice spisovné češtiny	7
1.1.3 Česká mluvnice Bohuslava Havránka a Aloise Jedličky	8
1.1.4 Mluvnice češtiny 2	9
1.1.5 Příruční mluvnice češtiny	10
1.1.6 Mluvnice současné češtiny	11
1.1.7 Akademická gramatika spisovné češtiny	13
1.1.8 Pravidla českého pravopisu	15
1.1.9 Historická mluvnice jazyka českého Jana Gebauera	17
1.2 Povaha morfonologické alternace	19
2 Korpusový výzkum	20
2.1 Metoda získání materiálu	20
2.1.1 Výběr substantiv deklinačního typu <i>žena</i> s dlouhým vokálem	20
2.1.2 Ruční výběr dvojslabičných apelativ	22
2.1.3 Výběr lemmat s nejednotnou kvantitou nebo kvalitou kořenného vokálu	24
2.1.4 Výsledky	30
2.2 Ověřování tvrzení z mluvnic a Pravidel	40
2.2.1 Vliv lexikálního významu na alternaci	41
2.2.2 Hlásková stavba jako faktor podmiňující alternaci	43
2.2.3 Odvozeniny od slov s alternujícím kořeným vokálem	44
2.2.4 Alternace kořenného vokálu v mluvených korpusech	46
2.3 Srovnání s AGSČ a MSC	47
Závěr	49
Literatura	50

Úvod

V bakalářské práci se budu zabývat alternací kořenného vokálu v některých femininech deklinačního typu *žena*. U některých feminin tohoto deklinačního typu dochází v určitých pádech ke krácení nebo změně kvality kořenného vokálu.

V první části se věnuji popisu zachycení tohoto jevu ve vybraných mluvnicích češtiny, různých vydáních Pravidel českého pravopisu a dalších publikacích. Výklady o alternaci kořenného vokálu feminin nejsou jednotné a liší se v některých aspektech, jako je např. názor na vliv lexikálního významu slov na alternaci jejich kořenného vokálu. Nezaměřuji se primárně na komparaci jednotlivých zdrojů, proto přehledy v mluvnicích podrobněji nesrovnávám. Takové srovnání by ukázalo zejména na vývoj pojetí tohoto jevu, případně na zaměření jednotlivých mluvnic a ostatních publikací. V této práci však chci především zachytit tento jev v úzu, tak, jak jej zaznamenává jazykový korpus.

V druhé části proto vytvořím seznam všech feminin deklinačního typu *žena*, u nichž dochází k alternaci kořenného vokálu. Tento seznam bude vytvořen na základě dat ze synchronního jazykového korpusu psané češtiny. Následně projde revizí, již provedu pomocí zpětného vyhledávání konkordančních řádků obsahujících tvary jednotlivých alternujících lemmat a posouzení adekvátnosti užití tvarů v jejich kontextech.

Část práce věnuji také ověřování platnosti některých tvrzení uvedených v literatuře a odrazu alternace v korpusech mluvené češtiny.

1 Krácení kořenného vokálu v dvojslabičných apelativech deklinačního typu *žena*

1.1 Zachycení jevu v literatuře

Tuto část práce věnuji souhrnu dosavadních poznatků a způsobů popisu zkoumané alternace. Z porovnání vybraných textů je znatelný i vliv přirozeného vývoje jazyka na jednotlivá zpracování tématu. Vybrané publikace, které o alternaci pojednávají, totiž zahrnují časové období více než 150 let, konkrétně roky 1860–2013¹.

Výklad o morfonologických alternacích dvojslabičných apelativ deklinačního typu *žena* je napříč mluvnicemi češtiny poměrně konzistentní a neliší se v žádných zásadních aspektech. Za cíl si texty obvykle kladou zachytit co nejvěrněji soudobý stav jazyka a za využití příkladů co nejpřesněji demonstrovat, za jakých okolností k alternaci dochází a jaké aspekty se na ní podílejí.

1.1.1 Česká mluvnice Jana S. Tomíčka

V *České mluvnici* Jana Slavomíra Tomíčka z roku 1860² (poprvé vyšla 1849) nalezneme poměrně zevrubný popis morfonologické alternace: „Dlouhé kmenové *á* krátí se v Gen. množného počtu v některých (jen dvojslabičných) jmenech: *jáma* – *jam*, v některých ale nikoli: *krása* – *krás*, což se i v písmě i v obyčejném životě děje. V písmě děje so [sic!] to též i v Instr. jednotného počtu, potom v Dat., Lok. a Instr. počtu množ., což se v obyčejném životě kolotá, přec ale více užívá se toho dle písma. V kterých jmenech kratí se tedy toto *á* v nadřčených pádech a v kterých nic?“³

Tomíček popisuje sémantický aspekt, který má podle něj na alternaci vliv: „V konkrétních jmenech (jimiž něco smyslového se udává), krátí se“ a udává jména, jichž se krácení *á* týká: *brána, pára, rána, tráva, vrána, sláma, žába, jáma, dráha, skála, kráva, řasa* (dnes jako *řasa*), *váha* atd. Pouze v genitivu se pak podle Tomíčka krátí *jáhla* – *jahel* a *játra* – *jater*⁴. Dlouhé *á* pak podle Tomíčka zůstává „ve statných jmenech

¹ Publikace jsem vybíral s ohledem na jejich závažnost pro českou normativní bohemistiku a doplnil je výběrem starších děl, která toto téma zpracovávají.

² Tomíček, J. S.: *Česká mluvnice. První (teoretický) díl*. Praha: Nákl. Bedřicha Tempského, 1860, s. 38.

³ U jednotlivých příkladů je vždy navíc uveden německojazyčný ekvivalent tištěný švabachem, do citací tyto překlady nejsou zařazeny.

⁴ V Jungmannově Slovníku česko-německém (1835) jsou *játra* uvedena jako plurál neutra vedle tvaru *játry*, která jsou plurálem feminina.

odtažných (abstraktních, jimiž něco nadsmyslového se udává)“ a uvádí příklady *krása, láska, spása, vláda, čáka* (naděje), *sláva, láska, váda, částka, ztráta, kára* (trest). O alternacích vokálu *í* pak pojednává takto: „Týž poměr (rovněž při dvouslabičných) nalézá se při proměně dlouhého *í* do *ě*, nejvíc jen ve třech následných jmenech: *míra, víra, díra*. [...] Toto *í* vzniklo ze starého *ě* (měra od měřiti): pročez původní *í* (z *ě* nevzniklé) mění se do *i* jen v následujících: *síla, žíla, síra*“.

Tomíček uvádí také výčet krátkých se slov obsahujících diftong *ou*: *moucha, trouba, houba, strouha, stoupa, kroupa, louka*. Uvádí také podmínku pro krácení u diftongu: „Následují-li po *ou* dvě souhlásky, nekrátí se“.

1.1.2 Brus jazyka českého

V Brusu jazyka českého („jejž sestavila kommisie širším sborem Matice České zřízená“) z roku 1877⁵ nalezneme v kapitole *Sklonění ženská* poznámky o zkoumané alternaci: „V instrum. jednot. čísla a genit., dat., lok. a instrum. množ. krátí se kmenová samohláska u konkrétních jmen dvouslabičných, která mají v nomin. před touto samohláskou jen jednu souhlásku“. Seznam konkrétních alternujících jmen zahrnuje slova: *brána, míra, máry, dráha, jáma, rána, sláma, víla, díra, síla*. Diftong *ou* se dle Brusu „z pravidla nekrátí kromě genit. množ. počtu: *houba, hub, moucha, much* atd.“ Stejně jako v Tomíčkově mluvnici se i zde objevuje zmínka o tom, že abstrakta krácení nepodléhají: „Abstraktní, jako: *krása, spása, sláva* atd., z pravidla nekrátí *á*, ač *víra* má věrou, věř, věřám atd.“

Ve třetím vydání Brusu z roku 1894, které je přepracované⁶, se objevuje podobný výklad. Chybí však informace o podílu sémantického aspektu na alternaci a výčet alternujících jmen je rozšířen o slova *lípa, hříva* nebo *hrouda*. Navíc je zde informace o vlivu dialektu na krácení: „co do krácení samohlásek působením dialektů pozorovati ve spisích od dávných dob mnohé kolísání“.

1.1.3 Trávníčkova Mluvnice spisovné češtiny

Mluvnice spisovné češtiny Františka Trávníčka z roku 1948⁷ uvádí, že ke krácení kořenného vokálu dochází „bud’ jen v 7. p. jedn. nebo jen v 2. p. mn. nebo v obou

⁵ *Brus jazyka českého*. Praha: nákl. Theodora Mourka, 1877, s. 6.

⁶ *Brus jazyka českého*. Praha: nakl. Jan Otto, 1894, s. 3.

⁷ Trávníček, F.: *Mluvnice spisovné češtiny*. Praha: Melantrich, 1948, s. 482–484.

těchto pádech a ještě v 3., 6. a 7. p. mn.“ Alternace zasahuje podle Trávníčka pouze slova dvojslabičná, která mají pouze jednu souhlásku před koncovkou (tedy s jednoslabičným tvarem genitivu plurálu). Je zde však i popsána výjimka: „z dvojslabičných jmen se souhláskovou skupinou a s 2. p. mn. dvojslabičným zkracují pomn. jména *jáhly* – *jahel* (jáhlám...) a *žábry* – *žaber* (žábrám...)“. Uvedeno je pět různých alternací kořeného vokálu: *á–a, í–i, í–ě, ů–o, ou–u*. Mezi slovy, která krácení podle Trávníčka nepodléhají, se objevují slova *šťáva* nebo *máry*.

Text je doplněn rozsáhlým seznamem feminin s alternujícím kořeným vokálem (prezentovaným jako výčet příkladů): *bába, bída, blána, bouda, brána, čára, dáma, díra, dláha, dráha, fůra, hlína, hlíza, houba, hrouda, chvála, jáma, jíva, kláda, kráva, kroupa, kůra, lípa, louka, míra, mísa, moucha, pára, rána, síla, skála, sláma, slíva, smlouva, strouha, tráva, trouba, váha, víla, víra, vláha, vrána, žába, žíla*.

U slov *váha* a *pára* jsou poznámky o vlivu ustálených spojení na alternace (*plnou parou* a *vlastní vahou*), u slov *dáma* a *žába* je uvedeno, že je alternace ovlivněna podobou označovaného (*dáma* se krátí, pouze pokud označuje životný denotát, *žába* se krátí v jiných tvarech, označuje-li zvíře, a v jiných tvarech, označuje-li mladou dívku).

1.1.3 Česká mluvnice Bohuslava Havránka a Aloise Jedličky

Ve vydání České mluvnice Bohuslava Havránka a Aloise Jedličky z roku 1980⁸ (dále ČM) je alternacím v deklinačním typu *žena* věnován sdílený prostor s alternacemi deklinačního typu *růže*. Dle ČM k alternacím dochází v „některých dvojslabičných [slovech] se samohláskou *á, í, ů, ou*“, a to především v genitivu plurálu. U některých z alternujících jmen dochází ke kolísání: „např. *jáhel* i *jahel*, *mís* i *mis*, *slív* i *sliv*, *fůr* i zast. *for*“. Zmíněny jsou také pády, v nichž mívají některá slova krátký kořený vokál, který však stojí vedle tvarů s vokálem dlouhým – dochází tedy ke kolísání. Toto se týká dativu plurálu, lokálu plurálu a instrumentálu plurálu. Jiná skupina slov je vymezena na základě přítomného kolísání i v instrumentálu singuláru. Alternace kořeného vokálu mají podle ČM v této skupině největší rozsah. Toto se snaží doložit příklady slov *rána* (v instrumentálu singuláru *ranou*) a *síla* (v instrumentálu singuláru *silou*). V ČM se objevuje odstavec o vázanosti krátkých tvarů na ustálená spojení,

⁸ Havránek, B. – Jedlička, A.: *Česká mluvnice*. Praha: Státní pedagogické nakladatelství, 1980, s. 159–161.

ovšem s konkrétními slovními významy: „U některých jmen je toto krácení vázáno jen na jistý význam nebo se vyskytuje jen v jistých ustálených spojeních; např. *váha* má v uvedených tvarech krátkou samohlásku jen ve významu ‚přístroj na vážení‘ (*vahou* i *váhou*,...), *míra* jen ve významu ‚stará plošná nebo dutá míra, měřice‘ (*měrou* i *mírou*,...); jinak je např. krátká samohláska v ustálených spojeních *tou měrou*, *plnou měrou*, *nepatrnou měrou*, podobně jako u jména *pára* je v 7. p. jedn. čísla krátké *a* jen ve spojení *plnou parou*.“ Podle ČM k alternaci vůbec nedochází u dvojslabičných slov s kořeným vokálem *y*, *é* a *ó* (u *ó* je zdůrazněn jeho výskyt v přejatých slovech) a u slov tříslabých nebo čtyřslabých i v tom případě, kdy se jedná o slova, která jsou s alternujícími dvojslabičnými slovy spojená nebo od nich odvozená (jsou uvedeny příklady *prasíla* – *prasílou* a *rovnováha* – *rovnováhou*).

1.1.4 Mluvnice češtiny 2

Poznámky týkající se krácení v dvojslabičných apelativech deklinačního typu *žena* nalezneme ve druhém dílu akademické *Mluvnice češtiny* (dále MČ 2).⁹ Najdeme je ve výkladu o deklinačním typu *žena*. MČ 2 vymezuje základní charakteristiku alternace jako změnu „kvantity, popř. zároveň kvality samohlásky základu“, která se projevuje u feminin převážně dvojslabičných, především v genitivu plurálu. Jedná se o pět konkrétních alternací: „*á/a skála* – *skal*, *í/i lípa* – *lip*, *í/ě míra* – *měr*, *ou/u houba* – *hub* a *ů/o*“. Dále uvádí, že střídání *ů/o* v současném jazyce ustupuje (tvary jako *for* a *kor* ustupují analogickým tvarům *fůr* a *kůr*) a hlásky *é*, *y* a *ó* nealternují vůbec. Ostatní pády, do kterých se dle MČ 2 alternace promítají, jsou dativ, lokál a instrumentál plurálu (tedy tvary s koncovkami tvořenými více než jednou hláskou). Samohláska naopak „nealternuje u jmen s tvarotvorným základem zakončeným skupinou souhlásek“ a „u jmen víceslabých, jako *zahrádka*, *katedrála* ap., ani u jmen odvozených předponou od těch dvouslabých jmen, u nichž k alternaci dochází, např. *nevíra*, *prasíla*, nebo s nimi složených, např. *slavobrána*, *rovnováha*.“ V textu je zmíněna i informace, že jsou změny kvantity a kvality samohlásek uvnitř tvarotvorného základu ve spisovné normě rozkolísány. O vývoji MČ 2 soudí, že ve všech tvarech kromě genitivu plurálu (jakožto tvaru s nulovou koncovkou) budou převažovat nekrácené podoby. Text obsahuje typograficky odlišenou poznámku týkající se uchovávání krácených forem ve frazémech apod.: „Tvary s alternací se

⁹ *Mluvnice češtiny 2*. Praha: Academia, 1986. s. 324–325.

udržují často jen v ustálených spojeních, v abstraktním užití ap. Např. *plnou parou, pustit žilou, podlehnout ranám osudu, být na vahách* ‚rozmýšlet se‘ atd. U ojedinělých jmen je s dvojí podobou tvaru spojen rozdíl významový, např. *brát se za někoho celou vahou svého postavení, podílet se významnou měrou*, ale pouze *vahou* ve významu ‚přístroj na vážení‘, *mírou* ve významu ‚rozměr, velikost‘.“ Seznam konkrétních slov, kterých se alternace týká, mluvnice neobsahuje.

Několik poznámek o alternacích v kmeni feminin se objevuje i v obecném výkladu o morfonologických alternacích (tamtéž, s. 278). V textu je uvedeno, že se alternace kmene (vedle deklinační koncovky) „podílejí na vyjadřování tvaroslovného významu a stávají se tak spoluformantem“, který zvýrazňuje funkci deklinačních koncovek. Autor zmiňuje, že někdy mají alternace i rozlišovací schopnost. U alternací kořenného vokálu ve dvojslabičných apelativech deklinačního typu *žena* by se jednalo o výše citovaná ustálená spojení a abstraktní užití.

1.1.5 Příruční mluvnice češtiny

Výklad o zkoumané alternaci v *Příruční mluvnici češtiny*¹⁰ (dále PMČ) se velmi podobá výkladu z MČ 2. Autorkou hesla je Zdenka Rusínová. Alternace se podle ní týká „řady dvojslabičných feminin majících v první slabice dlouhý vokál“ a dochází při ní ke krácení nebo změně vokálu v první slabice. Krácení zasahuje pouze vokály *á* a *í*, kvalita se mění z *í* na *ě* (*bída* – *běd*) a z diftongu *ou* na *u* (*moucha* – *much*). Přejatá slova se dle Rusínové obvykle nekrátí a ke krácení vůbec nedochází u vokálů *ó, é, ú* a *í*. Alternaci *ů–o* považuje za obvykle již neživou. Stejně jako v MČ 2 je i zde poznámka o nepřítomnosti krácení u slov se souhláskovou skupinou následující po dlouhém kořenném vokálu. Rusínová uvádí podmínku, že pokud ke krácení dochází, je tomu tak vždy v genitivu plurálu a dublety na této pozici jsou velmi řídké. Naopak o poměrně velkém množství kolísajících podob mluví v případě dativu, lokálu a instrumentálu. Větší měrou se krátké tvary udržují v ustálených spojeních. Jednotlivých variant se užívá k významovému nebo stylovému odlišení.

¹⁰ *Příruční mluvnice češtiny*. Praha: NLN, 1995. s. 256–257.

1.1.6 Mluvnice současné češtiny

Mluvnice současné češtiny z roku 2010¹¹ (dále MSČ) patří mezi novější mluvnice, které při výzkumu jazyka čerpají z dat obsažených v jazykových korpusech. Její specifická spočívá ve snaze reflektovat psaný i mluvený úzus a systematicky je porovnávat. Alternace kořenného vokálu dvojslabičných apelativ deklinačního typu *žena* je uvedena jako jedna ze tří hláskových alternací, které tento deklinační typ zasahují. Ke změně dochází dle MSČ v genitivu plurálu, popřípadě i instrumentálu singuláru, dativu plurálu, lokálu plurálu a instrumentálu plurálu. Ke změně naopak nedochází, je-li kmen zakončen souhláskovou skupinou.

Důležitým příspěvkem k výzkumu této alternace je tabulka popisující poměr výskytů krácených a nekrácených tvarů konkrétních slov (viz tabulka č. 1). Frekvence v tabulce nejsou uváděny číselně, ale na stupnici *nikdy – ojediněle – zřídka – někdy – stejně – často – zpravidla – vždy*.¹² Pod tabulkou je uvedena poznámka, že u některých slov pocházejí plurálové tvary převážně z víceslovných proprií, např. *K Lipám, V Lipách* apod. Je tedy zjevné, že Václav Cvrček, který je autorem tabulky, do výzkumu zahrnoval i propria. Mimo tabulku je uveden výčet slov, jejichž kořenné vokály jsou důsledně kráceny pouze v genitivu plurálu. K nim patří „*dáma – dam, smlouva – smluv, houba – hub, kroupa – krup, moucha – much, prakticky vždy bouda – bud, zpravidla díra – děr, trouba – trub, někdy u slov bába (zde je pro některé významy i v Nsg. varianta baba), mísa, slíva, víla a zřídka u slova š/čmouha*“.

Vliv ustálenosti na krácení kořenného vokálu je vyložen i s konkrétními příklady s uvedenými přibližnými poměry mezi krácenými a nekrácenými podobami: „V mnoha případech jde o volnou variaci, např. *podzemní dráhou/drahou* (poměr zhruba 3:1), *oběžnou dráhou/drahou* (2:1), *vodní párou/parou* (1:1). Někdy je ale mezi krátkým a dlouhým tvarem významový nebo stylový rozdíl (často využitý v rámci frazémů): je pouze *poznámka pod čarou, zasloužit se významnou měrou, ale vysokou mírou nezaměstnanosti/odpovědnosti/inflace* apod., *všeho s mírou*; je jen *plnou parou*, ale *(ohřát) nad párou*; podobně je *jednou ranou* (skoro vždy), ale *tržnou ranou/ránou* (1:1) atd.“

¹¹ Cvrček, V. a kol.: *Mluvnice současné češtiny*. Praha: Karolinum, 2010, s. 171–172.

¹² Tato stupnice užitá v tabulkách popisujících deklinační paradigmata odpovídá procentuálnímu zastoupení jednotlivých tvarů. Cvrček volí tuto nepřesnou škálu úmyslně, aby bylo možné data lépe srovnávat. *Nikdy* a *ojediněle* značí míru výskytu pod 1 %, *někdy* odpovídá 10–35 %, *stejně* 35–65 %, *často* 65–90 %, *zpravidla* 90–99 % a *vždy* nad 99 %. (MSČ 2010, s. 150)

Tabulka č. 1: Frekvence alternací kořenného vokálu v Mluvnici současné češtiny¹³

	Isg.	Gpl.	Dpl.	Lpl.	Ipl.
<i>bída</i>	ojed.	zpr.	(stejně)		
<i>blána</i>	někdy	zpr.	(stejně)		
<i>brána</i>	často	zpr.	zpr.		
<i>čára</i>	často	vždy	často		
<i>dráha</i>	někdy	vždy	často		
<i>hlína</i>	zř.	někdy	(zř.)		
<i>hrouda</i>	nikdy	zpr.	(někdy)		
<i>chvála</i>	nikdy	(zpr.)	zř.		
<i>jáma</i>	zř.	zpr.	(někdy)		
<i>kráva</i>	zř.	vždy	často		
<i>lípa</i>	stejně	zpr.	stejně*	stejně*	často
<i>míra</i>	stejně	vždy	zř.		
<i>pára</i>	stejně	vždy	zpr.		
<i>rána</i>	často	vždy	zpr.		
<i>síla</i>	zpr.	vždy	vždy	vždy	vždy
<i>skála</i>	někdy	vždy	zpr.		
<i>sláma</i>	zř.	—	—	—	—
<i>spára</i>	nikdy	nikdy	zř.		
<i>stoupa</i>	(zř.)	(někdy)	—	(nikdy*)	—
<i>strouha</i>	někdy	zpr.	—	stejně*	někdy
<i>šťáva</i>	nikdy	zř.	(zř.)	(zř.)	zř.
<i>tráva</i>	někdy	vždy	(často)	(často)	stejně
<i>váha</i>	stejně	vždy	(často)	zpr.	stejně
<i>víra</i>	ojed.	zpr.	(někdy)		
<i>vrána</i>	ojed.	vždy	(stejně)	(stejně)	(někdy)
<i>žába</i>	ojed.	vždy	zř.	(někdy)	zř.
<i>žíla</i>	zpr.	vždy	často	zpr.	často

¹³ Údaje v závorce jsou nejisté kvůli nedostatečnému počtu dat. Údaje označené hvězdičkou znamenají, že je poměr dlouhých a krátkých tvarů ovlivněn náležitostí jednoho z tvarů k víceslovným propriím (v MSC jsou uvedeny příklady *K Lipám*, *V Lipách*, *Na Stoupách*, *Ve Struhách*).

1.1.7 Akademická gramatika spisovné češtiny

Podle Akademické gramatiky spisovné češtiny z roku 2013¹⁴ (dále AGSČ), která své jazykové analýzy opírá stejně jako MSČ o data z jazykových korpusů, se kořený vokál *á* nebo *í* určitých dvojslabičných substantiv deklinačního typu *žena* krátí v některých případech důsledně, pravidelně a bezvýjimečně, jindy však dochází ke kolísání. Některá feminina s diftongem *ou* v kořeni mají v genitivu plurálu podobu s kořeným vokálem *u*.

V závislosti na krácení a nekrácení kořeného vokálu v jednotlivých pádech vymezuje AGSČ specifické podtypy deklinačního typu *žena*.

Pro vokál *á* je to šest podtypů: *zpráva*, *dáma*, *šťáva*, *vrána*, *rána*, *váha* (viz tabulka č. 2). Podtyp *zpráva* zahrnuje velkou většinu feminin deklinačního typu *žena*, jejichž kořený vokál se nekrátí. Jedná se o feminina více než dvojslabičná, z dvojslabičných to jsou slova s třemi souhláskami na začátku a s dvěma souhláskami, které následují po kořeném vokálu nebo předcházejí koncovce. Podtyp *dáma* zahrnuje slova, kde se krácení objevuje pouze v genitivu plurálu. Podtyp *šťáva* má v genitivu plurálu a instrumentálu plurálu dublety, přičemž dlouhé tvary jsou v tabulkách uvedeny na prvním místě, což znamená, že frekvenčně převládají. Podtyp *vrána* má krátký tvar vždy v genitivu plurálu a dublety (s častější dlouhou variantou) v dativu plurálu, lokálu plurálu a instrumentálu plurálu. V textu je do tohoto podtypu zařazeno slovo *žába*. Podtyp *rána* má dublety v instrumentálu singuláru, dativu plurálu, lokálu plurálu a instrumentálu plurálu. V těchto dubletách převažují krátké tvary. V genitivu plurálu se vyskytuje pouze krátký tvar. K tomuto podtypu jsou přiřazena slova *brána*, *dráha*, *skála*, *tráva*, *kráva*, *pára* a *jáma*. Posledním podtypem pro slova s kořeným vokálem *á* je podtyp *váha*, ke kterému AGSČ přiřazuje slova *blána* a *čára*. Tento podtyp je totožný s podtypem *rána* s jediným rozdílem. V genitivu plurálu se objevuje dubleta s převažující krátkou variantou. AGSČ také poukazuje na problémové zařazení slova *sláma*: „Problematické je z hlediska zachování či změn kvantity vokálu *-á-* femininum *sláma*, neboť jeho plurálové tvary nemáme (z pochopitelných důvodů) doloženy. V 1sg výrazně převládá zachování délky vokálu (*slámou*), avšak ojediněle se

¹⁴ Štícha, F. a kol.: *Akademická gramatika spisovné češtiny*. Praha: Academia, 2013, s. 339–341.

vyskytne i vokál zkrácený, a to i v kvalitním textu, jak ukazuje následující věta značně stylizovaného uměleckého stylu: [...] *a osvěcoval svým stříbrným světlem slamou posetý, před mýma očima rozprostřený dvůr*. Krátký vokál tu není třeba pokládat za chybu.“

Tabulka č. 2: Deklinační podtypy pro kořený vokál *á* podle AGSČ¹⁵

	Isg	Gpl	Dpl	Lpl	Ipl
<i>zpráva</i>	zprávou	zpráv	zprávám	zprávách	zprávami
<i>dáma</i>	dámou	dam	dámám	dámách	dámami
<i>šťáva</i>	šťávou	šťáv/šťav	šťávám	šťávách	šťávami/šťavami
<i>vrána</i>	vránou	vran	vránám/vranám	vránách/vranách	vránami/vranami
<i>rána</i>	ranou/ránou	ran	ranám/ránám	ranách/ránách	ranami/ránami
<i>váha</i>	vahou/váhou	vah/váh	vahám/váhám	vahách/váhách	vahami/váhami

Pro kořený vokál *í* rozlišuje AGSČ pět podtypů: *tíha*, *žíla*, *lípa*, *díra* a *víra* (viz tabulka č. 3). Všechna slova s nealternujícím kořeným vokálem náležejí k podtypu *tíha*. Tvary podtypu *žíla* kopírují podobu tvarů podtypu *rána*, mají tedy na všech potenciálně alternujících pozicích paradigmatu dublety s převažující krátkou podobou, kromě genitivu plurálu, kde je pouze krátký tvar. Podtyp *lípa* má ve všech uvedených tvarech dublety, v instrumentálu singuláru a instrumentálu plurálu převažují krátké tvary, v ostatních tvarech formy dlouhé. Podtypy *díra* a *víra* jsou vymezené pro kvalitativně alternující kořený vokál. Podtyp *díra* má ve všech tvarech kromě genitivu plurálu dublety s převažující podobou s vokálem *í*, v genitivu plurálu stojí krátký tvar s *ě*, zřídka se však objevuje i tvar *dír*. Tvary podtypu *víra* se téměř shodují s podtypem *díra*, pouze v instrumentálu singuláru je možný pouze tvar s vokálem *í*, v genitivu plurálu je pak možný pouze tvar s vokálem *ě*. AGSČ uvádí pouze dvě slova s alternujícím vokálem *í*, která jsou skloňována podle některého z uvedených podtypů a netvoří podtyp samostatný: slovo *síla* je skloňováno podle podtypu *žíla* a slovo *bída* patří k podtypu *víra*. K nealternujícímu lemmatu *tíha* patří „feminina s dvojicí či trojicí konsonantů před koncovkou nebo na začátku slova“.

¹⁵ Tabulky byla přeformátována pro lepší přehlednost ve formátu této práce. Tabulky publikované v AGSČ uvádějí ve sloupcích deklinační podtypy a v řádcích jednotlivé pády.

Tabulka č. 3: Deklinační podtypy pro kořený vokál *í* podle AGSČ

	Isg	Gpl	Dpl	Lpl	Ipl
tíha	tíhou	tíh	tíhám	tíhách	tíhami
žíla	žilou/žílou	žil	žilám/žílám	žilách/žílách	žilami/žílami
lípa	lipou/lípou	líp/lip	lípám/lipám	lípách/lipách	lipami/lípami
díra	dírou/děrou	děr (dír)	dírám/děram	dírách/děrách	dírami/děrami
víra	vírou	věr	vírám/věram	vírách/věrách	vírami/věrami

V rámci zachování systematického přístupu rozlišuje AGSČ dva deklinační podtypy i u slov s alternujícím diftongem *ou*: podtyp *houba*, kde se diftong *ou* v některých pádech mění na *u*, a podtyp *roura*, kde k alternaci nedochází v žádném pádě. K podtypu *houba* jsou přiřazena slova *bouda*, *hrouda*, *kroupa*, *louka*, *moucha* a *smlouva*. K nealternujícímu podtypu *roura* AGSČ řadí dvojslabičná feminina s dvojicí konsonantů před vokálem a tři konkrétní feminina *chlouba*, *zhouba*¹⁶ a *čmouha*. V AGSČ jsou obsaženy i poznámky o vlivu homonymie na kolísání tvarů *touh/tuh* a o ojedinělých výskytech tvaru *smlouv* (které však nepovažuje za chybné).

1.1.8 Pravidla českého pravopisu

Jednotlivá vydání Pravidel českého pravopisu přistupují k alternacím kořených vokálů dvojslabičných feminin deklinačního typu *žena* (resp. *ryba*) vesměs stejně. Podobně jako mluvnické se většina vydání Pravidel snaží co nejpřesněji zachytit podobu alternace a nastítné faktory, které ji ovlivňují.

Pravidla z roku 1902¹⁷ uvádějí: „Mnohá jména dvouslabičná vz. *ryba*, jež mají slabiku první dlouhou, zkracují ji v instr. jedn. a gen., dat., lok. a instr. množ.“ Tento úvod je doplněn příklady tvarů čtyř slov v uváděných pádech: *rána*, *brána*, *žíla* a *hrouda*. Poté následuje část, kde je popsáno kolísání u jednotlivých slov: „Při jiných jménech táž změna dílem jen částečně se děje, dílem nikdy.“ Důležitá je poznámka explicitně vyvracející vliv významu slov na krácení: „Výklad, že by se jména významu hmotného krátila a jiná nekrátila, je mylný.“ V tomto vydání Pravidel jsou také zmínky o tom, že

¹⁶ Autor textu uvádí, že náležitost slova *zhouba* k deklinačnímu podtypu *roura* je pouze hypotetické, a to z důvodu nedostatku dat.

¹⁷ *Pravidla hledící k českému pravopisu a tvarosloví*. Praha: Císařský královský školní knihosklad, 1902, s. 34.

„při některých jménech vyskytují se ve všech pádech tvary se samohláskou dlouhou a spolu také s krátkou“. Vedle sebe tak stály tvary *strouha* a *struha*, *sůva* a *sova*, *kůra* a *kora* apod.

V Pravidlech z roku 1913¹⁸ je alternace vokálu ve dvojslabičných femininech deklinačního typu *žena* zmíněna v části, kde se pojednává obecně o kvantitě a krácení. Značení kvantity v písmu je zde chápáno jako úzce spojené s výslovností: „Když kvantita je ve výslovnosti obecně zřetelná, bývá psaní snadné.“ Největší problémy jsou podle tohoto vydání Pravidel se zapisováním kvantity vokálu *i/í*, kde bývá výslovnost nezřetelná. Kvantitu vokálu by mělo být možné rozlišit na základě dvou pravidel: „Jest a) pravidlo, že za *e*, *ě* bývá u slov příbuzných *í*, a nikoli *i* (hřích, a nikoli hřích, poněvadž v příbuzném slově hřešiti jest -e- na tom místě [...]) a podle toho měl by býti rozk. způsob říd' (říd' proti ředitel). Ale je také b) ode dávna pravidlo pro některé rozk. způsoby, aby se kořenná jejich samohláska krátila, jak to viděti v *koupiti* – *kup* [...]; podle toho pravidla mělo by k říditi, stč. řediti, býti bud' starým krácením řed', nebo novým krácením řid'. Pravidla a) a b) si tu odporují, lze se tedy kolísati mezi psaním říd' a řid'.“ Jako jeden z příkladů je uveden genitiv plurálu slova *bříza*: „Podle pravidla a) má býti bříz (srov. březový atd.). Proti tomu jest opět pravidlo b), které žádá zkrácení, jako je v *houba* – *hub*, *brána* – *bran*, *síla* – *síl* atp.; krácením starým bylo by *břez*, ale to se nedochovalo; krácením novým mělo by býti *bříz*, ale tak se neříká. Ve výslovnosti nechává se -í- nezkrácené, a tak i píšeme.“ Pro slova, jejichž vokál se neznačí v souladu s těmito pravidly, Pravidla konstatují, že „není pro pravopis pomoci jiné, než pamatovati si, které psaní ovládlo, a toho se držeti“.

Pravidla z roku 1941¹⁹ popisují alternace v deklinačním typu *žena* společně s alternacemi v deklinačním typu *duše* (dnes *růže*). Seznamy²⁰ jmen s alternujícím vokálem jsou rozřazeny podle pádů. Slova, jejichž tvary se krátí pouze v genitivu plurálu, jsou *bába*, *bída*, *bouda*, *dáma*, *díra*, *fůra*, *houba*, *hrouda*, *chvála*, *jáhla*, *jáma*, *jíva*, *kroupa*, *mísa*, *moucha*, *sláma*, *slíva*, *smlouva*, *strouha*, *touha*, *trouba*, *víra*, *žába*. Slova krátící navíc i v dativu, lokálu a instrumentálu plurálu jsou *kráva*, *skála*, *tráva*, *žába*. Pod tímto seznamem je uvedena poznámka: „Znamená-li slovo *žába* tolik co

¹⁸ *Pravidla českého pravopisu*. Praha: Císařský královský školní knihosklad, 1913, s. 8–9.

¹⁹ *Pravidla českého pravopisu*. Praha: Školní nakladatelství pro Čechy a Moravu v Praze, 1941, s. 29–30.

²⁰ Pravidla z roku 1941 se snaží o zaznamenání všech alternujících dvojslabičných feminin deklinačních typů *žena* a *duše*, neprezentují je jako příklady.

„mladá dívka“, krátí jen v 2. p. množ.“ V instrumentálu singuláru krátí navíc *blána*, *brána*, *čára*, *dláha*, *dráha*, *kůra*, *lípa*, *míra*, *pára*, *rána*, *síla*, *váha*, *vrána*, *žila*. V seznamu je zdůrazněno, že víceslabičná slova odvozená od slov alternujících (*pablána* a *rovnováha*) kořenový vokál nekrátí. U slova *váha* Pravidla připouštějí dublety ve všech zmiňovaných pádech, *kůra* a *míra* mohou mít obě podoby v týchž pádech s výjimkou genitivu plurálu a u slov *brána* a *pára* je dubleta možná pouze v instrumentálu singuláru. Podobně jako u Havránka a Jedličky²¹ je zde poznámka o významových rozdílech mezi jednotlivými variantami: „*míra* krátí, znamená-li měřicí jednotku (=měřice), na př. *živnost o padesáti měrách*, ale nekrátí, znamená-li měřicí náčiní (=krejčovská míra a pod.); podobně *váha* (na př. *vážíme na váhách*, ale pojednání o měrách a vahách).“

Také v Pravidlech z roku 1957²² jsou alternace v deklinačním typu *žena* postaveny vedle alternací deklinačního typu *růže* (zde jako *nůše*). Výklad se v žádných podstatných rysech neliší od Pravidel z roku 1941. Rozšířen je výklad o vlivech způsobujících nekrácení. Podle tohoto vydání Pravidel se nekrátí „jména mající v základu *ý, é, ó*“, „jména mající před koncovkou skupinu souhlásek“, „jména s dlouhou předponou“, „jména trojslabičná a čtyřslabičná“ a „mnohá jména jiná“. Přibyla i poznámka o krácení slova *parou*, a to pouze v ustáleném spojení *plnou parou*.

V Pravidlech z roku 1993²³ je prostor věnovaný alternacím v deklinaci substantiv nejmenší. Objevuje se zde pouze informace, že „[v] kořenu některých tvarů téhož slova se střídají samohlásky krátké a dlouhé“ a následují příklady slov různých deklinačních typů.

1.1.9 Historická mluvnice jazyka českého Jana Gebauera

Ve třetím díle Gebauerovy Historické mluvnice jazyka českého, vzniklém v roce 1896²⁴, najdeme o zkoumané alternaci rozsáhlý výklad. Hodné pozoru je také pojednání o procesu dloužení, který zasahoval taktéž dvojslabičná feminina deklinačního typu *žena* (zde jako deklinační typ *ryba*). V něm se píše, že je-li tvar genitivu plurálu jednoslabičný, bývá dlouhý, přičemž v jiných pádech je táž slabika

²¹ Havránek, B. – Jedlička, A.: *Česká mluvnice*. Praha: Státní pedagogické nakladatelství, 1980, s. 159–161.

²² *Pravidla českého pravopisu*. Praha: Nakladatelství ČSAV, 1957, s. 41–43.

²³ *Pravidla českého pravopisu. Školní vydání*. Praha: Pansofia, 1993, s. 28.

²⁴ Gebauer, J.: *Historická mluvnice jazyka českého*. Díl III. Tvarosloví. Praha: Nakladatelství ČSAV, 1960, s. 182–186.

krátká. Jako příklad uvádí autor slova *strana* (*strán*), *pata* (*pát*), *hlava* (*hláv*), *straka* (*strák*), *noha* (*nóh*) apod.²⁵

O krácení kořenného vokálu deklinačního typu *ryba* Gebauer píše: „Mnohá substantiva, která mají sing. nom. dvouslabičný a v něm slabiku první dlouhou, zkracují ji v jazyku nynějším v sing. instr. a v plur. gen., dat., lok. a instr., tedy v těch pádech, kde proti nominativu dvouslabičnému je tvar jednoslabičný nebo trojslabičný, anebo kde v tvaru dvojslabičném je koncovka dlouhá [...] V textech starších poznáváme toto měnění kvantity jen potud, pokud se tam kvantita trochu bedlivě označuje, tedy od stol. XVI, poněkud také z doby starší.²⁶ Staré krácení bylo podle Gebauera přítomno kromě plurálu i v duálu. Rozdíl mezi starým a novým krácením Gebauer demonstruje na starých dokladech tvaru *hrozú* (instrumentál plurálu slova *hróza* – *hrůza*, který je zkrácením tvaru *hróza*) a novějších dokladech tvaru *hruzou* (který je zkrácením novějšího *hrůza*). Gebauer také připomíná možnou existenci dvou různých tvarů zároveň. Dubletní tvary podle něj vznikaly „novotvořením“, kdy vedle tvarů s „kmenovou slabikou právem dlouhou“, např. *strúha*, stály v jiných pádech tvary s „kmenovou slabikou právem krátkou“, např. instrumentál *struhou*. Analogií se pak utvořily dlouhé i krátké varianty pro ostatní pády. Gebauer uvádí i faktory ovlivňující krácení, mezi něž řadí i sémantický aspekt: „Substantiva, ve kterých krácení se děje, jsou většinou významu hmotného a mají v koncovce kmenové souhlásku jen jednu; ale pravidlem bez výjimky to není. Také jsou tu mnohé nestejnosti podle nářečí.“ Gebauer zmiňuje i tendence v krácení v soudobém jazyce: „vidí se, že pravidelnosti ubývá, zejména že jazyk pozdější v mnohých případech krácení nemá, kde v jazyku starším po právu bylo [...]; nejpevněji drží se bývalé zkrácení v plur. gen.“ V závěru textu je uveden poměrně rozsáhlý seznam dokladů feminin podléhajících krácení. Seznam obsahuje tato slova²⁷: *brána*, *brázda*, *dráha*, *chvála*, *jáhla*, *jáma*, *játry*, *kráva*, *máry*, *páry*, *rána*, *skála*, *sláma*, *sláva*, *tráva*, *váha*, *zpráva*, *žábra*, *čára*, *hlína*, *kníha*, *lípa*, *mísa*, *síla*, *síra*, *slíva*,

²⁵ Všechny tvary Gebauer dokládá odkazy na staročeskou literaturu.

²⁶ Krácení ve staročeském období Gebauer zdůvodňuje: „Že bylo toto měnění kvantity také v době staré, usuzujeme z analogie nč. a z příkladův, jako jsou *míra* – *měrou* atp.; kdyby krácení zde vykládané bylo z doby nové, měli bychom *míra* – *mírou* atd., jako je *síla* – *silou* atp.; jest však *měrou*, a v tom je důkaz, že krácení je z doby, kdy se ještě říkalo *miera* atd., t. j. z doby stč.“

²⁷ Pro potřeby práce jsem vypsál jen novočeské podoby slov (kromě podoby *kníha* s dlouhým vokálem).

žíla, bída, díra, míra, víra, hrůza, kůra, fůra, šňůra, bouda, hrouda, houba, kroupa, louka, moucha, mouka, smlouva, strouha, trouba.

Slovo *kníha* nebývá v novějších publikacích uváděno, protože jeho dlouhá podoba z psaného jazyka ustupuje během 19. století.²⁸

1.2 Povaha morfonologické alternace

Na základě poznatků z mluvnic češtiny a různých vydání Pravidel českého pravopisu o změnách kvantity (nebo kvality) zasahujících kořenný vokál dvojslabičných apelativ deklinačního typu *žena* jsem se rozhodl shrnout informace, ve kterých panuje napříč mluvnicemi a Pravidly shoda:

Alternace zasahuje především dvojslabičná slova, která mají v kořeni vokály *á* nebo *í* nebo diftong *ou*. U vokálu *á* dochází ke změně na *a*, u vokálu *í* ke změně kvantity na *i* nebo kvality na *ě*, u diftongu *ou* dochází ke změně na *u*. K alternaci docházelo v minulosti i u vokálu *ů*, který se měnil na *o*.

Alternace byly v minulosti důsledné, postupem času se objevují dubletní podoby. Nejčastěji ke změně kořenného vokálu dochází v genitivu plurálu, zasaženy bývají i plurálové tvary dativu, lokálu a instrumentálu a singulárové tvary instrumentálu. Krácené podoby nebo tvary se změněnou kvalitou vokálu se často zachovávají v ustálených slovních spojeních a frazémeh.

Alternace zpravidla nezasahuje slova, která mají po kořenném vokálu dva konsonanty a slova mající tři a více slabik.

²⁸ Podle nástroje SyD. Dostupné online na <www.syd.korpus.cz>.

2 Korpusový výzkum

Pro korpusový výzkum jsem využil materiálu obsaženého v korpusech *SYN2015* a *SYN*. *SYN2015* je synchronní korpus psané češtiny vytvořený kolektivem pracovníků Českého národního korpusu. Jeho velikost činí přibližně 100 milionů tokenů (120 milionů s interpunkcí, protože interpunkční znaménka tvoří samostatné tokeny). V současné době se jedná o nejnovější reprezentativní korpus tzv. řady *SYN*. Ta obsahuje synchronní korpusy psané češtiny: čtyři reprezentativní korpusy označené podle roku svého vzniku (*SYN2000*, *SYN2005*, *SYN2010* a *SYN2015*) a tři korpusy obsahující pouze texty psané publicistiky (*SYN2006PUB*, *SYN2009PUB* a *SYN2013PUB*). Šest z těchto korpusů (všechny kromě *SYN2015*) slučuje korpus *SYN*, který není referenční (jeho velikost se rozšiřuje postupným přidáváním nových verzí korpusů) ani reprezentativní (vzhledem k obsahu tří rozsáhlých korpusů složených výhradně z publicistických textů). I přesto je však cenným nástrojem pro výzkum jazyka, především díky své velikosti.

Konkrétní dotazy jsem zadával v dotazovacím jazyku *Corpus Query Language* (CQL) v korpusovém manažeru *KonText* dostupném přes internetový prohlížeč.²⁹

2.1 Metoda získání materiálu

Získávání první sady korpusových dat (pro výzkum monoftongů) jsem rozdělil do tří fází. V první fázi jsem z korpusu extrahoval apelativa náležící k deklinačnímu typu *žena*. Následně jsem z vygenerovaného seznamu ručně vybral dvojslabičná lemmata. Ve třetí, závěrečné fázi jsem je postupně po dávkách vkládal do CQL dotazů, pomocí kterých jsem mohl určit, zda se na některých pozicích paradigmatu kořenový vokál krátí. Vzniklý seznam lemmat jsem následně revidoval a vyřadil některá lemmata, u kterých se alternace neprojevují.

2.1.1 Výběr substantiv deklinačního typu *žena* s dlouhým vokálem

Nejprve bylo třeba získat z korpusu seznam femininních substantiv. Díky tomu, že jsou korpusy řady *SYN* morfologicky označovány a lemmatizovány, stačilo zadat

²⁹ Rozhraní *KonText* je dostupné z adresy <<http://kontext.korpus.cz>>.

správnou hodnotu pozičního atributu *tag*, a to „N“ na prvním místě v šestnáctimístném řetězci znaků (pro vyhledání substantiv) a „F“ na třetím místě (pro vyhledání feminin). Struktura dotazu by tedy mohla vypadat takto: *[tag="N.F.*"]*. Dotazovací jazyk CQL ovšem nabízí možnosti zjednodušení a urychlení práce. Proto jsem se do dotazu rozhodl zařadit i aspekty náležitosti lexému ke konkrétnímu deklinačnímu typu a přítomnosti dlouhého vokálu v jeho lemmatu. CQL umožňuje práci s logickými operátory, díky nimž lze mezi jednotlivými prvky dotazu realizovat logické operace. Já jsem využil tzv. logického součinu, který znamená současnou platnost několika výroků a je reprezentován „a“ nebo „a zároveň“. V CQL se „a zároveň“ značí pomocí ampersandu (&). Úsporná forma dotazu kombinujícího různé požadavky nemusí být žádoucí ve všech případech. Obzvláště pak, je-li např. při výzkumu určité množiny prvků podstatné analyzovat různé podmnožiny nebo poměry zkoumaných prvků v těchto podmnožinách obsažených; v tom případě je totiž důležité získávat dílčí poznatky i na mezistupních výzkumu. Rozhraní *KonText* pro tento proces postupné analýzy nabízí nástroje, jako je pozitivní a negativní filtrování výsledků dotazů. Kdybych využil těchto nástrojů, vyhledal bych nejdříve všechna substantiva, poté z nich vybral tokeny označované jako feminina, z nich dále vybíral jen substantiva deklinačního typu *žena* atd. Mezistupně, kterými bych musel projít, nám však nepřinášejí žádné cenné poznatky (pro tento výzkum například není podstatné, jak velká část všech feminin obsahuje dlouhý vokál), přestože mohou poukázat na případné chyby, jichž se v procesu výzkumu můžeme dopouštět.

Lemmata substantiv deklinačního typu *žena* jsou vždy zakončena formantem *-a*, v CQL dotazu jsem tedy specifikoval jejich podobu na „lemmata, jež tvoří nulový až nekonečný řetězec znaků, po němž následuje grafém *a*“. V CQL má pak dotaz podobu *[lemma="*.a"]*.

Lemma obsahující dlouhý vokál (jakékoli kvality) definuji jako „lemma, jež tvoří nulový až nekonečný řetězec znaků, po němž následuje jeden z dlouhých vokálů *á, é, í, ó, ú, ů* nebo *ý*, přičemž je nutné rozlišovat grafické podoby *ú* a *ů* (seznam je tvořen zapsáním jednotlivých znaků do hranatých závorek), po němž následuje nulový až nekonečný řetězec znaků“. V CQL se takové lemma značí *[lemma=".*[áíéóúůý].*"]*.

Kombinací dvou posledně zmíněných podmínek a patřičné morfologické značky se dají vyhledat lemmata, která jsou označkována jako substantiva a feminina, zároveň obsahují dlouhý vokál a zároveň jsou zakončena na *a*. Výsledný dotaz má tuto podobu:

```
[tag="N.F.*"&lemma=".*[áíéóúůý].*a"]
```

Při generování seznamu lemmat jsem postupoval po jednotlivých vokálech, abych se vyhnul nutnosti následného třídění. V korpusu *SYN2015* jsem hledal postupným zadáváním těchto sedmi dotazů:

```
[tag="N.F.*"&lemma=".*á.*a"]
```

```
[tag="N.F.*"&lemma=".*í.*a"]
```

```
[tag="N.F.*"&lemma=".*é.*a"]
```

```
[tag="N.F.*"&lemma=".*ó.*a"]
```

```
[tag="N.F.*"&lemma=".*ú.*a"]
```

```
[tag="N.F.*"&lemma=".*ů.*a"]
```

```
[tag="N.F.*"&lemma=".*ý.*a"]
```

Frekvence výskytů substantiv rozřazených podle jednotlivých vokálů, které obsahují, a součet jejich lemmat je uveden v tabulce č. 4. Můžeme si povšimnout nižší frekvence substantiv obsahujících vokál *ó*, což je způsobeno skutečností, že tento vokál se nachází převážně ve slovech nečeského původu.

Tabulka č. 4: Frekvence a počet lemmat substantiv deklinačního typu *žena* s dlouhým vokálem

vokál	á	í	é	ó	ú	ů	ý	celkem
frekvenc	74416	39084	11035	3401	6207	7234	15740	157121
e	6	8	6	9	8	7	5	9
lemmata	2351	1057	482	327	118	224	350	4909

2.1.2 Ruční výběr dvojslabičných apelativ

V druhé fázi jsem ze seznamů apelativ deklinačního typu *žena* s dlouhým vokálem vybíral dvojslabičné lexémy. Z výsledků každého ze sedmi výše uvedených dotazů jsem vytvořil frekvenční seznamy lemmat, které jsem následně exportoval do souboru formátu XLSX. V programu Microsoft Excel jsem pomocí funkce DÉLKA

seřadil lemmata podle délky řetězce znaků, načež jsem na jejím základě vyřadil příliš dlouhá lemmata a na základě frekvence v korpusu jsem vyřadil hapax legomena, tedy lemmata, která mají v korpusu jediný výskyt. Přestože by bylo možné použít algoritmus založený na selekci podle počtu samohlásek nebo slabikotvorných souhlásek, rozhodl jsem se pro čistě ruční výběr. Na jeho konci jsem dosáhl počtu dvojslabičných lemmat uvedeného v tabulce č. 5.

Tabulka č. 5: Dvojslabičná lemmata

vokál	á	í	é	ó	ú	ů	ý	celkem
lemmata	181	123	45	29	17	42	27	464

Lemmat obsahujících vokál *ú* značně ubylo, a to především z důvodu, že seznam obsahoval velké množství slov s prefixem *ú-*.

V seznamu stále zůstala propriální lemmata jako *Čína*, *Trója* nebo *Bára*, ale také např. slovo *Jáma*, které se vyskytuje v místních názvech. Propria nejsou předmětem mého výzkumu, proto jsem je odstranil. Odstraňovaná slova jsem identifikoval podle velkého písmena na začátku (apelativa jsou i na začátku větných úseků lemmatizována s malým písmenem). Stav po odebrání proprií a tedy finální materiál, s nímž jsem pracoval v další fázi, je uveden v tabulce č. 6.

Tabulka č. 6: Dvojslabičná apelativa

vokál	á	í	é	ó	ú	ů	ý	celkem
lemmata	141	108	34	23	13	38	24	381

Propria tvořila přibližně 18 % položek, po jejich odstranění jsem získal 381 různých dvojslabičných substantiv, jejichž základní slovníkový tvar obsahuje dlouhý vokál. V následující fázi se věnuji výběru těch z nich, kterých se týká krácení tohoto vokálu.

2.1.3 Výběr lemmat s nejednotnou kvantitou nebo kvalitou kořenného vokálu

2.1.3.1 Tvorba CQL dotazu

Ve třetí fázi jsem z 381 dvojslabičných lemmat vybíral ta, jichž se týká zkoumaná alternace vokalické kvantity nebo kvality. Zvolil jsem způsob, při kterém jsem data revidovaná v programu Microsoft Excel postupně vkládal pomocí specializovaných dotazů do korpusu a na základě jejich výsledku vytvořil finální seznam slov, u nichž dochází ke krácení kořenného vokálu, založený čistě na korpusových datech.

Jednotlivé CQL dotazy byly započaty morfologickou značkou pro substantiva a feminina [*tag="N.F.*"*], jejíž užití zamezuje chybám, které by mohly vzniknout v případě homonymie lemmat nebo slovních tvarů. Hlavní část dotazu tvoří samotné položky ze seznamu dvojslabičných lemmat a podmínka, která je klíčová pro určení lemmat s kráceným vokálem.

Položky ze seznamu byly zadávány jako hodnoty atributu *lemma*, např. [*lemma="síla"*]. Část dotazu identifikující lemmata, jichž se týká krácení, je podobná části dotazu použitého při vyhledávání lemmat s dlouhým vokálem. V té jsem hledal „lemmata, která jsou tvořena nulovým až nekonečným počtem znaků, dlouhým vokálem určené kvality, nulovým až nekonečným počtem znaků a grafémem *a*“. Pro identifikaci krácení jsem ale musel navíc zkoumat vztah lemmat a konkrétních realizací v korpusu. Na tomto vztahu je možné sledovat změny v hláskové stavbě. Díky dvojslabičné formě zkoumaných lemmat lze krácení kořenného vokálu určit poměrně snadno, a to především díky tomu, že kořenný vokál bývá kromě koncového *a* jediným vokálem ve slově. Proto je snadné jej s jistotou určit za pomoci regulárních výrazů využitých v zadávaných dotaz.³⁰ Dvojslabičný lexém deklinačního typu žena, jehož se týká alternace kořenného vokálu, jsem definoval jako „slovo, jež se alespoň v jedné ze svých realizací obsažené v korpusu liší v kvantitě prvního vokálu od začátku slova od lemmatu, které je mu přiřazeno“. Jedná se tedy o lexém, jehož lemma má dlouhý kořenný vokál, ale některá z jeho realizací (odlišná v pádu nebo čísle od lemmatu) dlouhý kořenný vokál nemá. Slovní tvar, který definuji jako hodnotu atributu *word*, který neobsahuje dlouhý vokál, zapisuji stejně jako ten, který dlouhý vokál má, a to

³⁰ Komplikovanější je postup při vyhledávání alternujících lemmat obsahujících kořenný vokál *á* (viz kap. 2.1.3.2.1).

pouze s tím rozdílem, že celou podmínku neguji. Toho jsem dosáhl přidáním vykřičníku před rovnítko. Podmínka pro slovo s vokálem *í* poté vypadá takto: `[word!=".*í.*"]`.

Dotazy, které jsem pro výzkum vytvářel, jsem kombinoval pomocí znaku „&“ pro logický součin jednotlivých prvků. Spojil jsem tři zmiňované podmínky, tedy morfologickou značku pro femininní substantiva, konkrétní podobu lemmatu a podobu slovního tvaru bez dlouhého vokálu. Zde je příklad takového dotazu:

```
[tag="N.F.*"&lemma="síla"&word!=".*í.*"]
```

Tento dotaz vyhledá v korpusu všechny tvary slova *síla*, které neobsahují grafém *í*, tedy např. *sil*.

Pozornost jsem musel věnovat ještě jednomu detailu. Atribut *word* je tzv. case-sensitive – rozlišuje velikost písmen. Je tedy podstatný rozdíl mezi vyhledáváním `[word="síla"]` a `[word="Síla"]` (na rozdíl od atributu *lemma*, u kterého na velikosti písmen nezáleží – tvar *ŽLÁZA* je lemmatizován jako *žláza*). Pro získání tzv. case-insensitive výsledků u atributu *word*, u kterých je ignorována velikost písmen, jsem přidal na začátek jeho hodnoty speciální řetězec znaků „(?i)“. Příklad kompletního dotazu umožňujícího identifikovat alternující dvojslabičná apelativa deklinačního typu *žena* pak vypadá následovně:

```
[tag="N.F.*"&lemma="síla"&word!="(?i).*.í.*"]
```

2.1.3.2 Postup práce a revize výstupu

Protože by bylo zadávání všech 381 dotazů odděleně neproduktivní, rozhodl jsem se zadávat dotazy obsahující několik lemmat zároveň. Toho jsem docílil nahrazením jednoho slova v hodnotě atributu *lemma* za seznam lemmat. Teoreticky by bylo možné pro každý vokál zadat pouze jeden dotaz obsahující všechna lemmata z příslušného seznamu, já jsem se však rozhodl zadávat dotazy po menších částech, které tvořilo přibližně deset lemmat. Seznam se v CQL uvádí v závorkách a jeho položky jsou od sebe oddělovány svislou čarou. Pro vyšší míru přehlednosti jsem zalamoval řádky dotazu:

```
[tag="N.F.*"&
(
lemma="bída"/
lemma="bríza"/
lemma="bříza"
)
&word!="(?i).*í.*"]
```

Po zadání tohoto dotazu se zobrazí výsledek tvořený všemi krátkými tvary lemmat ze zadaného seznamu. V případě dotazu z příkladu je to 15 výskytů slovního tvaru *běd*. Při zadání delšího seznamu či lemmat s vysokou frekvencí však bývá konkordančních řádků mnohem více, proto bylo nutné zobrazit frekvenční seznam lemmat, z něhož jsem opsal všechny jeho položky, tedy lemmata, u nichž dochází ke krácení kořenného vokálu. Po zadání dotazu z příkladu se ve frekvenčním seznamu lemmat objeví pouze lemma *bída*, pod které spadají všechny nalezené tokeny. Znamená to, že lemmata *bríza* ani *bříza* nereprezentuje v korpusu žádný slovní tvar, který by neobsahoval grafém *í* (např. tvary jako **brizami* nebo **břiz*).

V případech, kdy jsem si u některých lemmat zobrazených ve frekvenčních seznamech nebyl jistý jejich adekvátností, využíval jsem funkce filtrování pomocí pozitivního filtru, díky níž jsem mohl zobrazit konkordanční řádky výskytů krátkých tvarů konkrétního lemmatu a na základě kontextu pak posoudit, zda nedošlo k chybě v lemmatizaci nebo zda se nejedná o další specifické případy, kdy se kořenný vokál lemmatu a jeho slovního tvaru neshodují.

2.1.3.2.1 Specifika vyhledávání alternujících substantiv s kořenným vokálem á/a

Utvořené CQL dotazy, které vyhledávají za pomoci negativní podmínky ³¹ slova obsahující alternující kořenný vokál, fungují spolehlivě pro vokály *í*, *é*, *ó*, *ú* a *ů*. Pro vokál *á* však vhodné nejsou. Negativní podmínka, která je založena pouze na identifikaci hlásek (*word!=".*á.*"*), totiž vyřazuje z vyhledávání všechny slovní tvary,

³¹ Např. *[tag="N.F.*"&lemma="míra"&word!="(?i).*í.*"]*.

kteře obsahují grafém *á* (např. při zadání dotazu kombinujícího morfologickou značku, lemma *díra* a podmínku, že slovní tvar nesmí obsahovat grafém *í*, získáme výskyty všech alternujících tvarů ve všech pádech, tedy *děř*, *děřami*, *děřám* a *děřách*). Po zadání dotazu s podmínkou pro vokál *á* je třeba z vyhledávek odstranit i všechny tvary v dativu a instrumentálu plurálu, protože obě jejich koncovky *-ám* a *-ách* (MČ 2 1986, s. 325) obsahují grafém *á*.³² Pominutí tohoto jevu by způsobilo nevyrovnané podmínky pro vyhledávání slov s jednotlivými vokály, výrazně by ovlivnilo analýzu vyhledaných slovních tvarů a způsobilo, že by některá z alternujících lemmat vůbec nebylo možné objevit.³³ Dotazy pro vokál *á* jsem tak musel rozšířit o další podmínky, které tento jev způsobený podobou koncovek spolehlivě eliminují (zde příklad pro lemma *brána*):

```
[tag="N.F.*"&lemma="brána"&(word!="(?i).*á.*"/word="( ?i).*a.*ám"/word="( ?i).*a.*ách")]
```

Tento dotaz vyhledá všechny slovní tvary, které jsou substantiva a zároveň feminina, jsou přiřazené lemmatu *brána* a splňují alespoň jednu z těchto podmínek:

- 1) Jejich slovní tvar neobsahuje grafém *á*.
- 2) Jejich slovní tvar obsahuje grafém *a* a je zakončen *-ám*.
- 3) Jejich slovní tvar obsahuje grafém *a* a je zakončen *-ách*.

Dávkové vyhledávání alternujících lemmat obsahujících kořenový vokál *á* tak vypadalo např. takto:

```
[tag="N.F.*"&
(
lemma="bába"/
lemma="bárka"/
lemma="báza"
)
&(word!="(?i).*á.*"/word="( ?i).*a.*ám"/word="( ?i).*a.*ách")]
```

³² Jiné dlouhé vokály se v koncovkách tvarů deklinačního paradigmatu typu *žena* nenacházejí.

³³ To by se mohlo stát v případě, že by se všechny tvary daného lemmatu s krátkým kořenovým vokálem objevovaly v korpusu výhradně v dativu plurálu nebo lokálu plurálu. Negativní podmínka by všechny tyto tvary odfiltrovala a korpusový manažer by při vyhledávání nezobrazil žádný výsledek.

2.1.3.2.2 Výzkum pluralií tantum

V předešlé fázi výzkumu (viz kap. 2.1.3.2) jsem při hledání slov s alternujícím kořeným vokálem vycházel z prototypické hláskové podoby lemmatu náležícího k deklinačnímu typu *žena*, tedy řetězce znaků, jenž je zakončen koncovkou *-a*. Česká slovní zásoba ovšem obsahuje značné množství substantiv, která se užívají pouze v plurálových tvarech – tzv. pluralia tantum. Tato slova označují např. názvy částí těla, oděvů, nemocí, nástrojů a zařízení, svátků atd. (MČ 2 1986, s. 51–52) Vzhledem k faktu, že se jejich singulárové tvary v úzu prakticky nevyskytují, nejsou jejich lemmaty tvary nominativu singuláru, ale nominativu plurálu. Tvary nominativu plurálu slov deklinačního typu *žena* jsou zakončeny koncovkami *-y* nebo *-i*. (tamtéž, s. 325) Vytvořil jsem dva dotazy, každý pro jednu ze jmenovaných koncovek. Deklinační typ *žena* je mezi femininními deklinačními typy jediný, v jehož paradigmatu mají slova v nominativu plurálu koncovku *-y*, proto nebylo nutné využívat jiné podmínky pro vyhledávání, než je morfologická značka pro femininní substantiva a hlásková podoba lemmatu s obsaženým dlouhým vokálem, např.:

`[tag="N.F.*"&lemma="*i.*y"]`

Pro výskyty zobrazené v konkordančních řádcích jsem vytvořil frekvenční seznam lemmat, který jsem exportoval do formátu XLSX.

Vyhledávání pluralií tantum podle hláskového složení s koncovkou *-i* by mohlo být zavádějící. Tuto koncovku nominativu plurálu totiž sdílí paradigma deklinačního typu *žena* s paradigmatem deklinačního podtypu *kost*. (tamtéž, s. 330) Počet lemmat obsahujících dlouhý vokál uvnitř a grafém *i* na konci je však tak nízký, že nebylo nutné přistoupit v CQL dotazu k užší specifikaci. Frekvenční seznam lemmat jsem exportoval i v tomto případě. Poté jsem ručně vybral všechna dvojslabičná apelativa deklinačního typu *žena*. Do dotazů jsem je postupně vkládal stejným způsobem jako při vyhledávání alternací u lemmat zakončených *-a*, pro lemmata s kořeným vokálem *á* jsem v dotazu používal rozšířené podmínky.

Po zadání celkově čtyř dotazů pro vokály *á*, *í*, *ó* a *ů*³⁴ jsem v korpusu SYN2015 našel jediný výskyt plurale tantum *máry* s alternujícím kořeným vokálem: *v loubení s girlandami, kudy průvod nesl na marách znásilněnou noc*.³⁵

³⁴ Vokály *é*, *ý* a *ú* se nenacházely v žádném z nalezených dvojslabičných apelativních pluralií tantum deklinačního typu *žena*.

³⁵ Tento výskyt krátké podoby lemmatu *máry* je z beletristického textu Bohumila Hrabala.

2.1.3.2.3 Výzkum diftongů

Kromě dlouhých vokálů v první slabice dvojslabičných apelativ deklinačního typu *žena* dochází k hláskové alternaci u diftongu *ou*.³⁶ Ve své práci jsem se rozhodl alternace diftongů také analyzovat, proto jsem vytvořil dotaz kombinující tři podmínky, podobně jako u dotazů zadávaných za účelem výzkumu alternace monoftongů. První podmínkou je morfologická značka pro substantiva a feminina, druhá náležitost k lemmatu obsahujícímu zkoumaný diftong a zároveň zakončenému hláskou *a* a třetí podoba slovního tvaru, která zkoumaný diftong neobsahuje. Výsledkem byly tři dotazy pro tři různé diftongy, které se v češtině objevují: *ou*, *au* a *eu*.³⁷ Dotaz pro diftong *ou* vypadal následovně:

```
[tag="N.F.*"&lemma=".*ou.*a"&word!="(?i).*ou.*"]
```

Po zadání dotazů (do korpusu *SYN2015*) obsahujících diftongy *au* a *eu*, typické pro slova nečeského původu, nebyly nalezeny žádné výsledky. To znamená, že diftongů *au* a *eu* se kvalitativní alternace netýká.

Dotaz obsahující diftong *ou* vrátil 4 035 výskytů. Po zobrazení frekvenčního seznamu lemmat a vyřazení delších než dvojslabičných výrazů jsem získal seznam deseti lemmat, jejichž diftong *ou* obsažený v kořeni se neshoduje s vyskytujícími se slovními tvary. Tato lemmata jsou *houba*, *smlouva*, *louka*, *moucha*, *bouda*, *trouba*, *kroupa*, *stoupa*, *hrouda* a *chlouba*.

Nakonec jsem ještě vytvořil šest dotazů, jejichž účel byl identifikovat případná pluralia tantum obsahující v kořeni alternující diftong. Pro každý z diftongů *ou*, *au*, *eu* jsem vytvořil dva dotazy, pro tvrdou a měkkou koncovku. Výsledky na existenci pluralií tantum s alternujícím diftongem v kořeni nepoukázaly.

2.1.4 Výsledky

Po zadání dotazů obsahujících všechna lemmata se v následně generovaných frekvenčních seznamech objevilo celkem 37 lemmat, která vykazují slovní tvary s nestejným kořenným vokálem. Jsou jimi: *bába*, *bída*, *blána*, *blůza*, *brána*, *čára*, *dáma*, *díra*, *dráha*, *flóra*, *hlína*, *chvála*, *jáma*, *kóta*, *kráva*, *lípa*, *máma*, *míra*, *mísa*, *pára*, *píča*,

³⁶ Např. MČ 2, s. 325.

³⁷ V domácích slovech (citově neutrálních) je pouze dvojhláska *ou*, v slovech cizího původu nebo v domácích slovech emocionálně zabarvených také *au* a *eu*. (MČ 1, s. 34)

píka, rána, síla, skála, sláma, slíva, šťáva, tóga, tráva, váha, víra, vrána, žába, žába, žíla, žláza. Tato lemmata rozřazená podle obsaženého kořenného vokálu jsou uvedena v tabulce č. 7. Do tabulky je zařazeno i 10 lemmat s alternujícím diftongem.

Již při prvním zhlédnutí vygenerovaného seznamu je patrné, že je v rozporu s tím, jak je krácení kořenného vokálu popisováno v mluvnících češtiny.³⁸ Nejen proto je nutná následná revize všech výskytů lemmat v korpusu a vyhodnocení, zda není některé z lemmat v seznamu z jiného důvodu než vlivem zkoumané hláskové alternace.

Nejasností a chyb jsem si všímal již v průběhu vznikání seznamu, jejich odstraňování jsem ovšem odložil až do fáze jeho následné revize. Díky tomu, že jsem pominul průběžnou revizi, jsem mohl vytvořit seznam vycházející čistě z korpusových dat, zcela bez zásahu ovlivněného subjektivním jazykovým povědomím. Na takto „syrových“ výsledcích je možné lépe zhodnotit vhodnost užití metody výzkumu, případně i míru její chybovosti.

³⁸ V Mluvnici češtiny 2 je např. uvedeno, že nealternují hlásky *é, ý, ó*. (MČ 2, s. 325)

Tabulka č. 7: Alternující dvojslabičná apelativa deklinačního typu *žena*

á	í	ó	ů	ou
bába	bída	flóra	blůza	houba
blána	díra	kóta		smlouva
brána	hlína	tóga		louka
čára	lípa			moucha
dáma	míra			bouda
dráha	mísa			trouba
chvála	píča			kroupa
jáma	píka			stoupa
kráva	síla			hrouda
máma	slíva			chlouba
máry	víra			
pára	žíla			
rána				
skála				
sláma				
šťáva				
tráva				
váha				
vrána				
žába				
žábra				
žláza				

2.1.4.1 Chyby lemmatizace

Konkordanční řádky umožňující zobrazit kontext vyhledaných slovních tvarů poukazují na chyby v lemmatizaci, v některých případech kombinované i s chybami v morfoloickém značkování. V následujících případech uvádím některé z nich.³⁹

Lemma *brána*

Ve výsledcích dotazu na krácené podoby lemmatu *brána* se mezi správně lemmatizovanými tvary objevila propria, a to osobní jméno *Bran* a jméno rumunské obce *Bran*.

Lemma *čára*

Ačkoli je zařazení lemmatu *čára* do seznamu slov s alternující kvantitou (nebo kvalitou) kořenného vokálu namístě, protože korpus obsahuje značné množství genitivních a instrumentálních tvarů lexému *čára* v primárním významu ‚*souvislá řada bodů, linie*‘ (dle Slovníku spisovné češtiny ⁴⁰ – dále SSČ), objevuje se v konkordančních řádcích s krácenými tvary i tvar *čar*, který soudě dle kontextu náleží k lemmatu *čáry* ve významu ‚*kouzlo*‘ nebo ‚*čarování*‘⁴¹. Tento lexém je však v SSČ popsán pouze jako maskulinum s tvarem *čárů* v genitivu plurálu. Nejde zde však o chybu v morfoloickém značkování, jde pouze o doklady skloňování podle femininního vzoru, které je pravděpodobně způsobeno morfoloickou neprůhledností lexému (podle SSČ je *čáry* ve významu ‚*kouzlo*‘ plurale tantum), např. *čarodějnice, hodná své knihy čar*.

Lemma *dáma*

Ani lemma *dáma* není nutné ze seznamu vyjímát, protože např. v genitivu plurálu se důsledně krátí do tvaru *dam*. Do tohoto lemmatu jsou však zařazeny i slovní tvary, které k němu nenáleží. Jako *dáma* je lemmatizován nespisovný tvar slovesa *dát*, zde zřejmě dialektismus: *Dam ti ty papiry od auta a kliče a taky mam pro tebe nějake nove dyvka*. Dále se v konkordančních řádcích objevily i výskyty tvaru *dam* jakožto

³⁹ Počet konkordančních řádků s krácenými tvary všech revidovaných lemmat se pohybuje v řádu tisíců, vyhodnocování adekvátnosti kontextu každého z řádků by tedy bylo velmi pracné. Při práci s jazykovým korpusem počítám s přirozenou mírou chybovosti výsledků. Chyb, které jsou zvláště závažné a ovlivňují výsledky výzkumu, je možné si všimnout při pohledu na konkordanční řádky se zapnutou funkcí promíchání výsledků.

⁴⁰ *Slovník spisovné češtiny pro školu a veřejnost*. Praha: Academia, 1994.

⁴¹ Tamtéž.

citoslovce zpěvu: *Pam pa dam da dam dam*. Pod toto lemma spadá i proprium *DaM*, které je názvem architektonické kanceláře. Dále se tvar *Dam* objevil v *Macombs Dam Bridge*, který je anglickým výrazem pro přehradu, v chybně zapsaném názvu pařížské katedrály *Noter Dam* nebo jako část názvu *Thai Dam*, který označuje asijskou etnickou skupinu.

Lemma *chvála*

Pod lemma *chvála* bylo chybně zařazeno toponymum *Chvaly* (ve tvaru *Chval*).

Lemma *jáma*

K lemmatu *jáma* bylo přiřazeno anglické *jam*, a to jakožto část sousloví *jam session*, označujícího improvizaci setkání hudebníků.

Lemma *kráva*

Jako tvar lexému *kráva* je v korpusu označena část hebrejského *Krav maga* – názvu systému sebeobranných technik.

Lemma *máma*

V korpusu se nachází 2 545 výskytů slovních tvarů lemmatu *máma*, které neobsahují dlouhý vokál *á*. Všechny výskyty zastupují jediný tvar *mami*. SSČ však považuje *mami* za samostatný lexém, který je nesklonný a používá se „zvláště v oslovení“⁴². Stejný výklad nalezneme i ve Slovníku spisovného jazyka českého⁴³. Z tohoto důvodu jsem se rozhodl lemma *máma* ze seznamu vyjmout.

Lemma *pára*

Pod toto lemma spadají výrazy *par* z mnoha cizojazyčných pasáží, které jsou v korpusu obsaženy. Dále se objevilo i několik výskytů *par excellence* používaného v češtině. Chybně byl lemmatizován i termín *par* z oblasti golfu.

⁴² Tamtéž.

⁴³ *Slovník spisovného jazyka českého*, Praha: Nakladatelství ČSAV, 1960–1971. Dostupné v elektronické verzi z <<http://ssjc.ujc.cas.cz>>.

Lemma *rána*

Lemma *rána* zahrnuje zkratku *RAN* obsaženou v textu o informatice, objevil se i výskyt adjektiva *ranou*, u nějž došlo nejen ke špatné lemmatizaci, ale i chybnému morfologickému značkování: *těsně souvisí s ranou Bachovou snahou*.

Lemma *hlína*

Většina slovních tvarů lemmatu *hlína* nalezených v korpusu je osobním jménem *Hlin*, konkrétně galeristy Hlina Teta.

Lemma *lípa*

Slovní tvary spadající pod lemma *lípa* zahrnují i anglický výraz pro ret *lip*, dokonce se objevuje jeden výskyt slova *lip* jako citoslovce: *lip... lip... huráá*.

Lemma *mísa*

Pod lemma *mísa* spadá i několik iniciálových zkratk *MIS* a vlastní jméno postavy z beletristické knihy *Mis*.

Lemma *píka*

Lemma *píka* je pozoruhodné kvůli množství slovních tvarů odlišných lexémů, které byly jako tvary tohoto lemmatu označeny. Jedním z nich je *píka* označující barvu hracích karet: *na kterých kříže a piky vypadají úplně stejně*, dalším *piko* jakožto slangový výraz pro pervitin: *ty ubohý fetky, který tam takhle celý dny bloumaj a čekaj na piko*. Další ze slovních tvarů, *píka*, je druhovým jménem živočicha: *Pišťucha píka je chladnému horskému prostředí dokonale přizpůsobená*. Dalším tvarem je prefixoid *piko*, který se píše zpravidla dohromady se slovem, jemuž předchází, v některých vyskytujících se případech se ale píše odděleně: *kapesní (piko) LED projektor*. Tvar *piko* může sloužit i k zastření vulgarismu: *Ty piko, jó!* Objevuje se i citoslovečný tvar *pík*, zde značící zvuk hmyzu: *je v ní malý červíček a vyluzuje ty zvuky: pik pik pik* a tvar *piky*, využívaný při hře (zde aktualizované): *Piky piky na babu mám papíry na hlavu*. *Píka* se objevuje v popisu výslovnosti *pica*, výrazu označujícího psychickou poruchu. Ze všech 49 výskytů slovních tvarů bez vokálu *í* však není jediný nalezený tvar tvarem

lexému *píka* ve významu ‚zbraň podobná kopí‘⁴⁴. Proto jsem toto lemma ze seznamu vyřadil.

Lemma *síla*

K lemmatu *síla* byly chybně přiřazeny různé iniciálové zkratky *SIL*. Chybná lemmatizace a morfologické značkování se týká tvaru *sil* obsaženém v názvu písně *Sil jsem proso na souvrati*.

Lemma *víra*

Pod toto lemma spadá instrumentál vlastního jména *Věra*, tedy tvar *Věrou*. Zároveň se objevují i akuzativní tvary *věru*, které jsou pozůstatky starší podoby *věra* (ve smyslu ‚věření, důvěra, přesvědčení o něčem‘^{45 46}). Vzhledem k tomu, že se jedná o starší tvar, který je zakonzervován ve frazému *na mou věru*, nepovažuji jej za chybně lemmatizovaný, přestože je Slovníkem spisovné češtiny uváděn jako odlišné lemma archaického charakteru.

Lemma *žíla*

Lemma *žíla* zahrnuje tvar slovesa *žít* v názvu knihy *Žil jsem zbytečně*.

Lemma *flóra*

Tvar *Flora* lemmatizovaný jako *flóra* je v korpusu obsažen několikrát jako antroponymum (postava z beletristické knihy) a chrématonymum (pojmenování potravinářského výrobku nebo parfému). Tvary *flora* nebo *flor* se objevují i v cizojazyčných pasážích. K chybnému morfologickému značkování došlo u tvaru *flor* označujícího ‚jemnou průsvitnou látku‘ (SSČ).

Lemma *kóta*

Do lemmatu *kóta* se dostalo několik výskytů vlastních jmen *Kot* a *Koty*, dále výraz *kot* pro kocoura a slova z cizojazyčných textů v polštině nebo němčině. Tvar *Kota* se objevil v části názvu lesa *Kota Tinggi*. Ani jeden z 18 výskytů tvarů s krátkým vokálem *o* však nenáleží k lemmatu *kóta* ve významu ‚(v mapě) číslo udávající nadmořskou

⁴⁴ Dle SSČ.

⁴⁵ Tamtéž.

⁴⁶ Tvar *věra* je doložen v literárním díle Františka Ladislava Čelakovského nebo Josefa Václava Sládka. (Tamtéž)

výšku *zobrazeného bodu* (SSČ) nebo významu příbuzném. Proto jsem lemma *kóta* ze seznamu vyřadil.

Lemma *chlouba*

Jediný výskyt tvaru stejné podoby jako hypotetický tvar genitivu plurálu slova *chlouba*, tedy **chlub*, je imperativem slovesa *chlubit se*, což je zjevné z kontextu: *chlub se mandelinkám*.

Lemma *louka*

Na jevy spojené s lemmatizací tvarů slova *louka* poukazuje např. Česká mluvnice Havránek a Jedličky: „Jméno *louka* má v množ. čísle nejen *luk*, *lukám*, ale i 1. a 4. p. *luka* vedle *louky*, *loukám*,...” (Havránek–Jedlička 1981, s. 160) Porovnáme-li frekvenční seznamy slovních tvarů po vyhledávání podle lemmat *louka* a *luka*, zjistíme, že jsou všechny krátké tvary jako *lukách*, *lukami* nebo *lukám* lemmatizovány jako tvary slova *luka*, což je plurale tantum středního rodu (SSČ, SSJČ). Jako krátký tvar lemmatu *louka* byl lemmatizován pouze tvar *luk*, tedy genitiv plurálu. Některé výskyty tvaru *luk* jsou přiřazeny také lemmatu *luka*, přičemž při pohledu do konkordančních řádků není jisté, podle čeho bylo lemma totožných tvarů určeno (srov. např. tvar lemmatu *louka* v kontextu *typická jarní rostlina vlhkých luk a slatinišť* a tvar lemmatu *luka* v kontextu *druh, který ale roste často v podrostu vlhkých luk, kde světelný počitek není příliš vysoký*).

Lemma *stoupa*

Ze seznamu lemmat s kořeny obsahujícími diftong *ou* bylo neproblematictější lemma *stoupa*. Pod něj totiž spadají dva různé, zcela odlišné lexémy. Prvním z nich je výraz *stup* (maskulinum), který se používá v oblasti horolezectví a znamená ‚místo, kam se klade noha‘ (SSJČ): *ze všech sil kope špičkou boty, aby si udělal stup*. Druhým je výraz *stúpa* pocházející ze sanskrtu, jehož význam uvádí Akademický slovník cizích slov jako ‚buddhistická náb. stavba budovaná jako mohyla se schránkou na ostatky n. jako památník, dágoba‘⁴⁷. Doklady jako *stup* mohou být následkem nedůsledného psaní diakritiky. Mohou to ale také být případy kvantitativní alternace kořenného vokálu *stúpa* – *stup*. Krátké tvary, potenciálně alternující jako ve výskytu *jezdíme na exkurze*

⁴⁷ Akademický slovník cizích slov. Praha: Academia, 2001.

do hrodek a pyramid a pohřebišť a stup se při výzkumu monoftongických vokálů nemohly zobrazit, protože byly chybně lematizovány jako *stoupa*.

2.1.4.2 Různé varianty zápisu týchž lemat

Ukázalo se, že slovní tvary obsažené v korpusu, které nesdílí podobu kořenného vokálu se svým lemmatem, nejsou jen tvary ovlivněné morfonologickou alternací, kterou zkoumám, a tvary s chybnou lematizací nebo nesprávně přiřazenou morfológickou značkou, ale také tvary, jako je *flora*, které se v kořenném vokálu od svého lemmatu liší kvůli odlišnému grafickému zaznamenání.⁴⁸

První díl akademické mluvnice češtiny uvádí, že „[p]očešťování grafiky přejatých slov se v značném rozsahu projevuje označováním délky u samohlásek ve shodě s výslovností.“ (MČ 1 1986, s. 168)⁴⁹ Značení kvantity ovšem bylo podstatně ovlivněno kodifikací pravopisu, která byla měněna buď na základě snahy o přiblížení pravopisu výslovnosti, nebo snahy o sjednocení výslovnosti na základě určených pravopisných pravidel (Sedláček 1993, s. 137)⁵⁰. Ve vývoji psaní byly podstatné zásahy Pravidel českého pravopisu z roku 1957, poměrně striktně určující kodifikované tvary, nebo revize zpracování tohoto tématu v Pravidlech českého pravopisu z roku 1993, kdy byla některá pravidla změněna nebo byly připuštěny pravopisné dublety.

Problém s dvojím způsobem zápisu slov přejatých z cizích jazyků se týká tří slov ze seznamu exportovaného z korpusu. Ani jedno ze slov *flóra*, *tóga* a *blůza* nemá domácí původ. U slova *blůza* se píše *ů* a toto slovo je tak uváděno i ve slovnících, přestože systémově by jako slovo cizího původu mělo být psáno s *ú*. Slovo *flóra* se na české území dostalo z latinského jména bohyně rostlinstva *Flóry*, *tóga* z latinského *toga* a *blůza* pochází z německého *Bluse* (Rejzek 2015). Ve vydání Českého etymologického slovníku Jiřího Rejzka z roku 2015 se v heslech objevuje datace prvních dokladů daných slov v češtině, z níž jsem zjistil, že první dochované doklady jsou datovány do

⁴⁸ O způsobech a možnostech morfológické anotace variant viz např. článek *Slovní varianty a morfológická anotace korpusů* (Hlaváčová, J., In: *Grammar & Corpora, Gramatika a korpus 2007*. Praha: Academia, 2008, s. 161–168.

⁴⁹ Srov. slova *demokracie* (původně *démo-*), *etický* (z *éthos*) nebo *maturita* (z *mátúritás*). (MČ 1 1986, s. 104)

⁵⁰ Sedláček, M.: K vývoji českého pravopisu. Část 2. In *Naše řeč*, r. 76, č. 3, 1993. s. 126–138.

19. století. To znamená, že jsou zřejmě všechny tři tyto lexémy přítomny v české slovní zásobě relativně krátce.

V dokladech s krátkými vokály, jako je *ochrany fauny a flory celé planety*, není možné rozlišit, zda se jedná o zkrácené podoby dlouhých tvarů (*flory* jako genitiv singuláru slova *flóra*), nebo o tvary slov, která mají v nominativu singuláru krátký kořenný vokál (*flory* jako genitiv singuláru slova *flora*). Krácení kořenných vokálů *ó* a *ú*, respektive *ů* v těchto slovech tak nelze na základě výskytů z korpusu spolehlivě dokázat.

Jiným případem je lemma *píča*. Jakožto vulgarismus se toto slovo vymyká spisovnému jazyku, proto není jeho psaná podoba zakonzervována kodifikací a různé podoby psané formy odrážejí různé varianty výslovnosti. Krátká varianta *piča* (která se vedle dlouhé podoby užívá⁵¹) tak může být přímým odrazem krátké výslovnosti typické např. pro slezskou nářeční oblast. Ač existují doklady, které by na krácení kořenného vokálu mohly poukazovat (např. v *Žilinské ulici kopal do kola a nadával do cyklistických pič*), není možné s jistotou určit, zda k němu opravdu dochází.⁵²

2.1.4.3 Výsledky po revizi

Po revizi výsledků, při které jsem odstranil všechna lemmata, u nichž nelze s jistotou určit, zda podléhají morfonologickým alternacím, jsem došel k počtu 39 alternujících lemmat. Jejich seznam rozřazený podle konkrétních morfonologických alternací je uveden v tabulce č. 8.

⁵¹ Internetová jazyková příručka. Dostupné z <<http://prirucka.ujc.cas.cz>>.

⁵² Teoretickou možností by bylo zkoumání celého textu, v němž se krácený tvar vyskytl. Zadáme-li do rozhraní KonText dotaz na výskyty všech tvarů obou variant lexému (zde jsou obě varianty lemmatizovány v podobě s dlouhým kořenným vokálem) a omezíme vyhledávání podmínkou *within* pouze na konkrétní dílo, v tomto případě dotazem *[lemma="píča"]within <doc id="piuss_laskajesli" />*, zjistíme, že text obsahuje tvar s krátkým kořenným vokálem v nominativu singuláru. Kdyby byl autor textu (resp. jeho překladatel) v psaní varianty konzistentní, znamenalo by to, že uvedený příklad v *Žilinské ulici kopal do kola a nadával do cyklistických pič* na krácení kořenného vokálu nepoukazuje.

Tabulka č. 8: Morfonologické alternace dvojslabičných apelativ deklinačního typu *žena* – výsledky po revizi

á/a	í/i	í/ě	ou/u
bába	hlína	bída	bouda
blána	lípa	díra	houba
brána	mísa	míra	hrouda
čára	síla	víra	kroupa
dáma	slíva		louka
dráha	žíla		moucha
chvála			smlouva
jáma			trouba
kráva			
máry			
pára			
rána			
skála			
sláma			
šťáva			
tráva			
váha			
vrána			
žába			
žábra			
žláza			

2.2 Ověřování tvrzení z *mluvnic a Pravidel*

V závěrečné části práce se věnuji srovnání různých zachycení změny kvantity (nebo kvality) kořeného vokálu v dvojslabičných femininech deklinačního typu *žena* v literatuře (viz kapitola 1.1) se zachycením tohoto jevu v jazykových korpusech. Při porovnání budu pracovat s materiálem získaným v předešlé části práce (viz kapitola 2.1).

2.2.1 Vliv lexikálního významu na alternaci

2.2.1.1 Krácení u konkrét a abstrakt

V Tomíčkově mluvnici z roku 1860⁵³ nalezneme tvrzení, že je krácení kořenného vokálu (konkrétně vokálu *á*) ovlivněno lexikálním významem slov. Ke krácení by mělo docházet u konkrét, u abstrakt by pak měl kořenný vokál zůstat ve všech pádech stejný. Tato myšlenka⁵⁴ je přítomna i v Gebauerově historické mluvnici z roku 1896.⁵⁵ Gebauer však tvrdí, že se tento jev netýká všech slov a nejedná se tedy o bezvýjimečné pravidlo. Pravidla z roku 1902⁵⁶ domněnku o vlivu konkrétnosti a abstraktnosti na alternaci kořenného vokálu považují za mylnou.

Podle České lexikologie Čermáka a Filipce „[k]onkrétní substantiva označují substance jako nositele vlastností a dějů, smysly vnímatelné a spojené s názornými představami: lidi, živočichy, rostliny, části těla, zařízení, nástroje, výrobky ap. Abstraktní substantiva označují vlastnosti, děje a stavy pojaté jako samostatné substance: *ctnost, krása, červeň, let, chvála, spánek*“^{57 58}. Už mezi samotnými příklady abstrakt z této publikace se nachází lexém *chvála*, jehož zkrácené tvary obsahuje korpus SYN2015⁵⁹ a který byl proto zařazen do seznamu krátkých se substantiv.

Přestože je podle mluvnic směr vývoje krácení jasně popsán, rozhodl jsem se ověřit existenci krátkých tvarů slova *chvála* v diachronním korpusu. Využíval jsem korpusu *Diakon* přístupného v aplikaci SyD⁶⁰ (znamenající *synchronní a diachronní analýza*⁶¹), která původně slouží k všestrannému výzkumu variant v korpusech Českého národního korpusu.⁶² Tvary slova *chvála* v dnes primárním významu ‚*zhodnocení a uznání dobrých vlastností n. činů, projev uznání někomu, uspokojení s něčím*‘ (SSJČ) jsou

⁵³ Tomíček, J. S.: *Česká mluvnice. První (teoretický) díl*. Praha: Nákl. Bedřicha Tempského, 1860, s. 38.

⁵⁴ V souvislosti s krácením všech vokálů.

⁵⁵ Gebauer, J.: *Historická mluvnice jazyka českého*. Díl III. Tvarosloví. Praha: Nakladatelství ČSAV, 1960, s. 182–186.

⁵⁶ *Pravidla hledící k českému pravopisu a tvarosloví*. Praha: Císařský královský školní knihosklad, 1902, s. 34.

⁵⁷ Čermák s Filipcem zde čerpají z publikace Vladimíra Šmilauera *Nauka o českém jazyku* (1972).

⁵⁸ Čermák, F. – Filipce, J.: *Česká lexikologie*. Praha: Academia, 1985, s. 97.

⁵⁹ Ve výskytech *chorál díků, jásoť a chval, ačkoliv jsem byl u samých počátků chval a uctívání* apod.

⁶⁰ Přes rozhraní *KonText* je dostupný pouze diachronní korpus *Diakorp*, který obsahuje menší počet tokenů než *Diakon*.

⁶¹ Zdroj dostupný online na <<http://wiki.korpus.cz/doku.php/manual:syd>>.

⁶² Cvrček, V. – Vondříčka, P.: Výzkum variability v korpusech češtiny. In: Čermák, F. (ed.): *Korpusová lingvistika Praha*. Praha: NLN, 2011, s. 184–195.

doloženy např. v roce 1812: *Držím zákoníctví za chval nejhodnější* nebo v roce 1822: *a předce na ní chval a han do tisíce*.⁶³

Řada abstrakt se pak nachází mezi alternujícími slovy s kořeným vokálem í: *síla, víra, míra*⁶⁴, *bída*.

Domněnku, že je kategorie⁶⁵ konkrétnosti nebo abstraktnosti lexému podstatným faktorem ovlivňujícím morfonologickou alternaci kořeného vokálu dvojslabičných feminin deklinačního typu *žena*, považuji za mylnou z důvodu značného výskytu abstrakt, jež alternace zasahuje.

2.2.1.2 Krácení u polysémních lexémů

Podle některých mluvnic (např. Havránkovy–Jedličkovy nebo Trávníčkovy) je alternace kořeného vokálu ovlivněna lexikálním významem slova. Podoba deklinačního paradigmatu může mít specifickou podobu pro jednotlivé lexikální významy polysémního lexému.

Slovo *váha* se podle Pravidel českého pravopisu z roku 1941⁶⁶ v lokálu plurálu nekrátí, jestliže označuje „*přístroj na vážení*“ (SSJČ), pokud ale označuje „*vlastnost tělesa působenou zemskou přitažlivostí*“ (SSJČ), kořený vokál se mění na krátký. Uvedený příklad z těchto Pravidel zní: „vážíme na váhách, *ale* pojednání o měřácích a vahách.“ V Havránkově a Jedličkově České mluvnici⁶⁷ má však *váha* „krátkou samohlásku jen ve významu „*přístroj na vážení*“ (*vahou* i *váhou*,...)“. Tato tvrzení si protiřečí.

Krátké tvary slova *váha* se v korpusu SYN2015 objevují nezřídka v obou těchto významech. Význam „*přístroj na vážení*“ zastupují výskyty jako např. *Přirovnávala život k dvojramenným vahám, na nichž se jazýček nikdy nezastaví uprostřed; některé typy digitálních vah zvládají i roli hodin* a velké množství krátkých tvarů ve spojeních

⁶³ V dokladech z 19. století (a starších) převládají výskyty slova *chvála* ve významu „*druhá část kněžských a řádových hodinek; laudy*“ v SSJČ uváděném jako sekundárním. I zde je však možné nalézt zkrácené tvary, např. *Kdež zněl předtím božských chval zpěv líbezně* (1812).

⁶⁴ V abstraktním významu.

⁶⁵ Podle V. Mathesia je možné přisuzovat těmto třídám slov kategoriální charakter. (Čermák–Filipec 1985, s. 97)

⁶⁶ *Pravidla českého pravopisu*. Praha: Školní nakladatelství pro Čechy a Moravu v Praze, 1941, s. 29–30.

⁶⁷ Havránek, B. – Jedlička, A.: *Česká mluvnice*. Praha: Státní pedagogické nakladatelství, 1980, s. 159–161.

jazýček vah nebo *misky vah*. Po zadání dotazu *[word="misky"]**[word="v[aá]h"]* do korpusu *SYN2015* jsem zjistil, že všechny výskyty spojení *misky vah/váh* mají kořenný vokál slova *váha* krátký (46 výskytů)⁶⁸. Význam ‚*vlastnost tělesa působenou zemskou přitažlivostí*‘ se objevuje ve výskytech jako např. *pasivní nahromadění látek s vyšší molekulovou vahou*; *Lisujte šest hodin vahou 23 kg* nebo ve velkém množství spojení jako *plnou vahou*, *vlastní vahou*, *celou vahou* apod. Tvrzení z Pravidel českého pravopisu z roku 1941 nepovažuji v současném jazyce zachyceném korpusem *SYN2015* za platné.

2.2.2 Hlásková stavba jako faktor podmiňující alternaci

Jako jedna z podmínek ovlivňujících náchylnost k alternaci je v mluvnicích (např. AGSČ) a Pravidlech (např. vydání z roku 1957) zmiňována podoba hláskové stavby slov. Nejčastěji je uváděno, že k alternaci kořeného vokálu nedochází, nachází-li se za kořenným vokálem, resp. před koncovkou dva nebo více konsonantů.

Analyzoval jsem hláskovou stavbu slov obsažených v seznamu vygenerovaném korpusem. Výsledky uvádím v tabulce č. 9 (vokály jsou označeny písmenem V, konsonanty písmenem C):

Tabulka č. 9: Hlásková stavba lemmat alternujících dvojslabičných feminin deklinačního typu *žena*

Hlásková podoba	Počet
CVCV	21
CCVCV	16
CCCVCV	1
CVCCV	1
Celkem	39

Z tabulky vyplývá, že 38 z celkových 39 alternujících feminin má hláskovou podobu v souladu s tím, co uvádějí mluvnické, jedno slovo, *žábra*, má však mezi kořenným vokálem a koncovkou dva konsonanty. Toto slovo má v korpuse *SYN2015* 18 výskytů

⁶⁸ Po zadání dotazu do rozsáhlejšího korpusu *SYN* je to 1 064 výskytů tvarů s krátkým vokálem, stále však ani jeden výskyt tvaru *váh* (vyhledáváno 1. 8. 2016).

krátkého tvaru genitivu plurálu *žaber*⁶⁹. Tvzení Akademické gramatiky spisovné češtiny, že se substantiva s dvojicí konsonantů před koncovkou řadí k nealternujícímu podtypu *zpráva*, je tak nepřesné. Podobné informace o absenci alternace kořenného vokálu dvojslabičných feminin deklinačního typu *žena* se skupinou souhlásek mezi tímto vokálem a koncovkou nalezneme i v Mluvnici současné češtiny.

Etymologický slovník Jiřího Rejzka⁷⁰ uvádí, že slovo *žábry* bylo do českého jazyka přejato až Janem Svatoplukem Preslem v 19. století (z této doby přirozeně pocházejí i první dochované doklady užití tohoto slova v češtině). Tvary s krátkým kořenným vokálem jsou tak zřejmě tvořeny analogií.

2.2.3 Odvozeniny od slov s alternujícím kořenným vokálem

Podle mluvnic (např. České mluvnice Havránka a Jedličky⁷¹ nebo Mluvnice češtiny 2⁷²) nedochází k alternaci kořenného vokálu u feminin majících více než dvě slabiky, a to i v případech, jsou-li tato slova odvozena od alternujících. Toto tvrzení jsem se rozhodl ověřit hromadným zadáním 39 dotazů, tedy pro každé z alternujících lemmat, jež jsem extrahoval z korpusu. Abych vyhledal prefigovaná slova a kompozita, vycházel jsem z předpokladu, že každému slovu ze seznamu alternujících feminin musí předcházet nenulový řetězec znaků, např.:

```
[(  
lemma"..*bába"/  
lemma="..*brána"/  
lemma="..*blána"  
)]
```

Z frekvenčního seznamu lemmat vygenerovaného pro výsledky dotazu⁷³ jsem ručně vybral všechna slova, která jsou odvozena od alternujících slov ze seznamu

⁶⁹ V korpusu SYN je to 125 výskytů tvaru *žaber* (vyhledáváno 1. 8. 2016).

⁷⁰ Rejzek, J.: *Český etymologický slovník*. Voznice: LEDA, 2015.

⁷¹ Havránek, B. – Jedlička, A.: *Česká mluvnice*. Praha: Státní pedagogické nakladatelství, 1980, s. 159–161.

⁷² *Mluvnice češtiny 2*. Praha: Academia, 1986, s. 324–325.

⁷³ Frekvenční seznam obsahoval 171 lemmat, velkou část tvořila propria (např. *Elvíra*, *Vladimíra*) a slova obsahující stejné hlásky, ale neodvozená od slov s alternujícím kořenným vokálem (např. *membrána*, *kalamáry*, *zhouba*).

(viz 2.1.4.3) a mají v korpusu alespoň dva výskyty. Získal jsem 32 lemmat: *autodráha*, *biošťáva*, *časomíra*, *dvojbrána*, *gardedáma*, *hajzlbába*, *jednobrána*, *jetelotráva*, *kuličkodráha*, *nadváha*, *nerovnováha*, *nevíra*, *nitroblána*, *oposmlouva*⁷⁴, *podváha*, *prabába*, *praprabába*, *protisíla*, *protiváha*, *přemíra*, *rovnováha*, *rychlodráha*, *samochvála*, *sebechvála*, *siločára*, *slavobrána*, *supersíla*, *světočára*, *turbodíra*, *tvrdohouba*, *veledíra*, *vodováha*.

U každého z těchto slov jsem ověřoval, jestli se v korpusu SYN2015 nenacházejí jejich tvary s alternujícím kořeným vokálem. To jsem prováděl vyhledáváním všech tvarů jednotlivých lemmat a následným zobrazením frekvenčního seznamu slovních tvarů. Doklady alternace kořeného vokálu jsem našel u šesti kompozit nebo prefigovaných slov: *protiváha*, *rychlodráha*, *prasíla*, *tvrdohouba* a *siločára*.

V korpusu se objevil tvar genitivu plurálu *protivah* v tomto kontextu: *jež existují v charakteristickém rámci brzd a protivah*. Korpus SYN2015 obsahuje tři výskyty tvaru *protiváh* a pouze jeden výskyt tvaru *protivah*.

Tvar *rychlodrah* se v korpusu vyskytuje také pouze jednou: *Na rozdíl od předchozích rychlodrah jde totiž o výhradně čínský projekt*. Varianta s dlouhým vokálem se však v tomto pádě nevyskytuje vůbec.

K lemmatu *prasíla* jsou přiřazeny tvary *prasilami* a *prasil*: *Tak může vypadat pán nad prasilami života; pramenů a podzemních jeskyň jako sídel přírodních prasil*. Uvedené výskyty jsou jedinými doklady krátkých tvarů v tomto korpusu, varianty s dlouhým vokálem se však nevyskytují vůbec.

V korpusu se objevil i jeden výskyt tvaru *tvrdohub*, náležící k lemmatu *tvrdohouba*: *V některých případech (zejm. u tvrdohub) se mohou konidie tvořit ve stejných plodnicích*. Tvar **tvrdohoub* není doložen.

Nejvíce výskytů krátkých tvarů má slovo *siločára*. Zejména v genitivu plurálu frekvenčně převládá tvar s krátkým vokálem nad tvarem s vokálem dlouhým. V tabulce č. 10 uvádím výskyty tvarů slova *siločára* v korpusu ve všech pádech, v kterých se projevuje alternace kořeného vokálu u feminin deklinačního typu *žena*. Číslo v závorce značí počet výskytů.

⁷⁴ Výraz pro *opoziční smlouvu*.

Tabulka č. 10: Výskyty tvarů vybraných pádů slova *siločára* v korpusu *SYN2015*

Pád	Tvary
Isg	— (0)
Gpl	siločar (104)/siločár (2)
Dpl	siločarám (1)/siločárám (1)
Lpl	— (0)
Ipl	siločarami (9)/siločárami (4)

Z výsledků je zjevné, že se kořenový vokál slova *siločára* krátí téměř pravidelně v genitivu plurálu, krátké tvary se objevují i v dativu a instrumentálu plurálu.

Slovo *siločára* je důkazem, že slova odvozená od dvojslabičných feminin deklinačního typu *žena* s alternujícím kořenovým vokálem, která mají více než dvě slabiky, mohou v některých tvarech kořenový vokál krátit.⁷⁵

2.2.4 Alternace kořenového vokálu v mluvených korpusech

Možnosti zkoumání alternace kořenového vokálu dvojslabičných apelativ deklinačního typu *žena* v mluvených korpusech jsou zatím značně omezené. Jedním z dílčích problémů může být nespolehlivost distinkce vokalické kvantity vyslovené mluvčími⁷⁶, hlavním faktorem ovlivňujícím možnosti výzkumu je však, jak se ukázalo, množství dat v korpusu obsažených.

Pro výzkum v korpusu mluveného jazyka *ORAL2013* jsem vybral ze seznamu alternujících lemmat (viz 2.1.4.3) několik slov s vysokou frekvencí, konkrétně slova *váha*, *čára*, *síla* a *víra* a vyhledal jejich tvary v pádech s alternačním potenciálem⁷⁷ v podobách s krátkým i dlouhým vokálem. Tokeny obsažené v korpusu *ORAL2013* neprošly lemmatizací a chybí jim morfologické značkování, jednotlivé tvary jsem tak musel zadávat ručně⁷⁸. CQL dotaz pro slovo *váha* tak vypadal následovně:

⁷⁵ Tvar genitivu plurálu slova *čára* se vyskytuje v korpusu *SYN2015* častěji v podobě *čar* (412 výskytů), zřídka v podobě *čár* (2 výskyty). Distribuce je tak podobná jako u slova *siločára*.

⁷⁶ Výrazně spolehlivější by byla distinkce variant kvalitativních alternací, příp. rozlišování variant s kořenovými vokály *i* a *í*, u nichž bývá v artikulaci výrazný kvalitativní rozdíl.

⁷⁷ Instrumentál singuláru a genitiv, dativ, lokál a instrumentál plurálu.

⁷⁸ Na webu Českého národního korpusu je sice nástroj nazvaný *Vyhledávač variant v korpusech řady ORAL* (dostupný z <http://wiki.korpus.cz/doku.php/kurz:hledani_v_mluvenych_korpusech>), který po zadání lemmatu zobrazí tvary a jejich různé varianty, jejich souhrn však (k 4. 8. 2016) neobsahuje všechny varianty kořenového vokálu u alternujících feminin.

[word="vahou/váhou/vah/váh/vahám/váhám/váhách/vahách/vahami/váhami"]

Tato metoda je problematická kvůli homonymii, která může ovlivňovat výsledky u některých tvarů (v tomto případě může být *váhám* i tvar slovesa *váhat*).

V korpusu *ORAL2013* se nacházejí tři výskyty tvaru *vahou* a taktéž tři výskyty tvaru *váhou*. Výskyty dotazu pro slovo *čára* zahrnuje pouze 6 výskytů singulárního tvaru *čarou* a jediný výskyt dlouhé varianty *čárou*. Slovo *síla* v tomto korpusu zastupují krátké tvary instrumentálu singuláru (7)⁷⁹, genitivu plurálu (5) a instrumentálu plurálu (2). Dlouhé tvary se v tomto korpusu vůbec nevyskytují. Pro slovo *víra* byl nalezen jediný výskyt, a to dlouhého tvaru *vírou*.

Výsledky vyhledávek nelze spolehlivě interpretovat, svědčí pouze o přítomnosti alternujících tvarů v mluveném jazyce. Možnosti mluvených korpusů nejsou pro zkoumání této alternace dostačující. V teoretickém případě mnohonásobně většího počtu dat by bylo možné zkoumat např. vliv sociálních faktorů na tento jev nebo i rozdíly mezi jednotlivými nářečními oblastmi.

2.3 Srovnání s AGSČ a MSČ

Akademická gramatika spisovné češtiny a Mluvnice současné češtiny jsou mluvnice, které při svých výzkumech vycházejí z dat poskytnutých jazykovými korpusy. V obou mluvnicích jsou nastíněny základní faktory ovlivňující krácení u feminin deklinačního typu *žena*, jako je počet slabik lemmatu (ke krácení nedochází u tříslabičných a delších slov) nebo počet konsonantů mezi kořeným vokálem a koncovkou (ke krácení nedochází, je-li mezi kořeným vokálem a koncovkou více než jeden konsonant). Vůbec v nich však nejsou zmíněny případy, které se těmto pravidlům vymykají, třebaže se v úzu zachyceném korpusem vyskytují.

Popsány nejsou případy krácení ve slovech s více než dvěma slabikami, jako je *siločára*. Zmíněna nejsou ani alternující slova *žábřa* a *jáhla*⁸⁰, tedy slova obsahující dva konsonanty mezi kořeným vokálem a koncovkou. Plurale tantum *máry* není

⁷⁹ V závorkách je uveden počet výskytů.

⁸⁰ Slovo *jáhla* není zahrnuto v seznamu alternujících feminin (viz 2.1.4.3), protože jsem vycházel z dat korpusu *SYN2015*, kde se krátké tvary tohoto slova nevyskytují.

obsaženo v seznamu alternujících feminin v MSČ ani není zařazeno do deklinačních podtypů v AGSČ.

Ačkoli jsou výsledky výzkumu ovlivněny verzí korpusu, z níž jsou data čerpána, v korpusu *SYN2005*, s níž už autoři těchto mluvnic měli možnost pracovat, jsou všechny tyto případy krácení doloženy⁸¹.

Výčty jednotlivých feminin s alternujícím kořenným vokálem v AGSČ a MSČ se neshodují ani mezi sebou, ani se seznamem vycházejícím z korpusu *SYN2015* (viz 2.1.4.3). V AGSČ nejsou přítomna slova *bába, hlína, chvála, máry, míra, mísa, slíva, trouba, žábra* a *žláza*. V MSČ nejsou uvedena slova *bouda, dáma, díra, houba, kroupa, louka, máry, mísa, moucha, slíva, smlouva, trouba, žábra* a *žláza*. Nad rámec seznamu uvedeného v kap. 2.1.4.3 uvádí MSČ slova *spára, stoupa* a *strouha*.

⁸¹ V korpusu *SYN2005* jsou výskyty krátkých tvarů slov *jáhla, žábra, máry* a *siločára*.

Závěr

V této práci jsem se věnoval výzkumu změny kvantity nebo kvality kořenného vokálu některých dvojslabičných apelativ deklinačního typu *žena*.

V první části práce jsem popsal zachycení tohoto jevu v mluvnicích češtiny, vydáních Pravidel českého pravopisu a dalších publikacích. Ukázalo se, že výklady o alternaci nejsou ve všech publikacích jednotné a některá tvrzení v nich obsažená jsou dokonce protichůdná (např. rozdíly v alternaci kořenného vokálu u jednotlivých významů polysémních lexémů).

Druhá část práce obsahovala samotný korpusový výzkum alternace. V jednotlivých fázích jsem popsal metodu extrakce dat z jazykového korpusu a způsob vytváření CQL dotazů. Pozornost jsem věnoval i specifikům vyplývajícím z užití metody, která se týkala vyhledávání alternujících lemmat s kořenným vokálem *a*. Výstupem korpusového výzkumu je seznam dvojslabičných feminin deklinačního typu *žena*, jejichž tvary mají v některých pádech zkrácený nebo kvalitativně odlišný kořenný vokál. Do seznamu jsem kromě lemmat obsahujících vokály *a* a *i* zařadil také lemmata obsahující diftong *ou*, který se v některých pádech mění na *u*. Během práce s korpusem jsem sledoval nepřesnosti v lemmatizaci jednotlivých tvarů a popsal je. V rámci výzkumu jsem analyzoval i pluralia tantum.

V části korpusového výzkumu jsem ověřoval některá tvrzení uvedená v mluvnicích a dalších publikacích. Konkrétně jsem se zabýval vlivem lexikálního významu a hláskové stavby slov na alternaci jejich kořenných vokálů a odvozenin. Vybraná alternující slova jsem vyhledával i v korpusu mluvené češtiny, účelné porovnání s psanými korpusy však nebylo možné uskutečnit kvůli nedostatku dat.

Vzhledem k omezeným možnostem výzkumu alternace v mluvených korpusech by bylo vhodnější zkoumat tento jev v mluveném jazyce experimentálně. Poté by bylo možné zaměřit se na sociologické faktory nebo vliv dialektů na alternaci.

Literatura

- Akademický slovník cizích slov*. Praha: Academia, 2001.
- Brus jazyka českého*. Praha: nákl. Theodora Mourka, 1877.
- Brus jazyka českého*. Praha: nakl. Jan Otto, 1894.
- CVRČEK, V. – VONDŘIČKA, P.: *SyD – Korpusový průzkum variant*. Praha: FF UK, 2011. Dostupné na <<http://syd.korpus.cz>>.
- CVRČEK, V. a kol.: *Mluvnice současné češtiny*. Praha: Karolinum, 2010.
- ČERMÁK, F. – FILIPEC, J.: *Česká lexikologie*. Praha: Academia, 1985.
- Český národní korpus – *ORAL2013*. Praha: Ústav Českého národního korpusu FF UK, 2013. Dostupné na <<http://www.korpus.cz>>.
- Český národní korpus – *SYN*. Praha: Ústav Českého národního korpusu FF UK. Dostupné na <<http://www.korpus.cz>>.
- Český národní korpus – *SYN2005*. Praha: Ústav Českého národního korpusu FF UK, 2005. Dostupné na <<http://www.korpus.cz>>.
- Český národní korpus – *SYN2015*. Praha: Ústav Českého národního korpusu FF UK, 2010. Dostupné na <<http://www.korpus.cz>>.
- DOKULIL, M.: *Mluvnice češtiny. 1, Fonetika, fonologie, morfonologie a morfemika, tvoření slov*. Praha: Academia, 1986.
- GEBAUER, J.: *Historická mluvnice jazyka českého. Díl III. Tvarosloví*. Praha: Nakladatelství ČSAV, 1960.
- HAVRÁNEK, B. – JEDLIČKA, A.: *Česká mluvnice*. Praha: Státní pedagogické nakladatelství, 1980.
- Internetová jazyková příručka*. Dostupné na <<http://prirucka.ujc.cas.cz>>.
- JUNGSMANN, J.: *Slownjik německo-český*. Praha: Academia, 1989.
- KOŘENSKÝ, J. – KOMÁREK, M. a kol.: *Mluvnice češtiny. 2, Tvarosloví*. Praha: Academia, 1986.
- Pravidla českého pravopisu*. Praha: Císařský královský školní knihosklad, 1913.
- Pravidla českého pravopisu*. Praha: Nakladatelství ČSAV, 1957.

Pravidla českého pravopisu. Praha: Školní nakladatelství pro Čechy a Moravu v Praze, 1941.

Pravidla českého pravopisu. Školní vydání. Praha: Pansofia, 1993.

Pravidla hledící k českému pravopisu a tvarosloví. Praha: Císařský královský školní knihosklad, 1902.

KARLÍK P. – NEKULA, M. – RUSÍNOVÁ, Z. (eds.): *Příruční mluvnice češtiny*. Praha: NLN, 1995.

REJZEK, J.: *Český etymologický slovník*. Voznice: LEDA, 2015.

SEDLÁČEK, M.: K vývoji českého pravopisu. Část 2. In *Naše řeč*, r. 76, č. 3, 1993. s. 126–138.

Slovník spisovné češtiny pro školu a veřejnost. Praha: Academia, 1994.

Slovník spisovného jazyka českého, Praha: Nakladatelství ČSAV, 1960–1971.
Dostupné na <<http://ssjc.ujc.cas.cz>>.

ŠTÍCHA, F. a kol.: *Akademická gramatika spisovné češtiny*. Praha: Academia, 2013.

TOMÍČEK, J. S.: *Česká mluvnice. První (teoretický) díl*. Praha: Nákl. Bedřicha Tempského, 1860.

TRÁVNÍČEK, F.: *Mluvnice spisovné češtiny*. Praha: Melantrich, 1948.